

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 18:57:27 ; Search time 134 Seconds
(without alignments)
386.916 Million cell updates/sec

Title: US-09-638-693A-36_COPY_16_133
Perfect score: 616
Sequence: 1 ACMSADLEVTSTWVLLGV.....VIEPIVTNNQKLEAFMHKH 118

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	616	100.0	133	2	AAR63288	Aar63288 Polypepti
2	608	98.7	133	2	AAR63289	Aar63289 Polypepti
3	608	98.7	133	2	AAR63290	Aar63290 Polypepti
4	608	98.7	209	2	AAR63390	Aar63390 HCV poly
5	603	97.9	128	2	AAR37937	Aar37937 HCV NS4 r
6	601	97.6	127	2	AAR37936	Aar37936 HCV NS4 r
7	601	97.6	127	2	AAR37933	Aar37933 HCV NS4 r
8	599	97.2	128	2	AAR37932	Aar37932 HCV NS4 r
9	597	96.9	3021	9	ADK40820	Adk40820 HCV poly
10	597	96.9	3021	9	ADK40823	Adk40823 HCV poly
11	597	96.9	3021	9	ADK40824	Adk40824 HCV poly
12	590	95.8	133	2	AAR63287	Aar63287 Polypepti
13	590	95.8	133	2	AAR63286	Aar63286 Polypepti
14	579	94.0	3021	9	ADK40813	Adk40813 HCV poly
15	545	88.5	117	2	AAR37934	Aar37934 HCV NS4 r
16	545	88.5	117	2	AAR37935	Aar37935 HCV NS4 r
17	527	85.6	3023	2	AAR94462	Aar94462 Hepatitis
18	527	85.6	3023	9	ADK40803	Adk40803 HCV poly
19	481	78.1	3019	9	ADK40810	Adk40810 HCV poly
20	467	75.8	3002	7	ADM24822	Adm24822 Hepatitis
21	467	75.8	3011	2	AAW77397	Aaw77397 Hepatitis
22	467	75.8	3011	6	ABP71460	Abp71460 Amino aci
23	467	75.8	3011	8	ADH79949	Adh79949 E2 HCV en
24	467	75.8	3011	8	ADL72983	Adl72983 Hepatitis

25	467	75.8	3012	5	AAU99289	Aau99289 Hepatitis
26	467	75.8	3012	6	ABU61848	Abu61848 HCV H77 c
27	465	75.5	252	8	ADT77851	Adt77851 Hepatitis
28	464	75.3	3011	2	AAR66995	Aar66995 Hepatitis
29	464	75.3	3011	9	ADX40785	Adx40785 HCV poly
30	463	75.2	3011	5	AAU84597	Aau84597 HCV poly
31	462	75.0	237	1	AAP90138	Aap90138 Peptide e
32	462	75.0	363	2	AAR23999	Aar23999 Open read
33	462	75.0	363	2	AAR90933	Aar90933 HCV anti
34	462	75.0	382	1	AAP92048	Aap92048 Carboxy-t
35	462	75.0	382	1	AAP90182	Aap90182 C terminu
36	462	75.0	460	1	AAP92024	Aap92024 Polypepti
37	462	75.0	460	1	AAP90141	Aap90141 Protein s
38	462	75.0	592	2	AAR33565	Aar33565 CKS-HCV a
39	462	75.0	592	4	ABG69023	Abg69023 HCV recom
40	462	75.0	592	7	ABW01888	Abw01888 HCV CKS-C
41	462	75.0	594	2	AAR33566	Aar33566 CKS-HCV a
42	462	75.0	594	4	ABG69024	Abg69024 HCV recom
43	462	75.0	594	7	ABW01889	Abw01889 HCV CKS-C
44	462	75.0	597	2	AAR21571	Aar21571 HCV CKS-C
45	462	75.0	597	2	AAR33638	Aar33638 HCV C100D

ALIGNMENTS

RESULT 1
AAR63288
ID AAR63288 standard; protein; 133 AA.
XX AC AAR63288;
XX DT 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX DE Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
XX KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW Classification; immunisation; prophylaxis; serotyping.
XX OS Hepatitis C virus type 3.
XX PN WO9425601-A2.
XX PD 10-NOV-1994.
XX PF 27-APR-1994; 94WO-EP001323.
XX PR 27-APR-1993; 93EP-00401099.
XX PA 05-AUG-1993; 93EP-00402019.
XX PI (INNO-) INNOGENETICS NV SA.
XX Maertens G, Stuyver L;
XX WPI; 1994-358277/44.
XX N-PSDB; AAQ78040.
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX Claim 11; Page 125; 404pp; English.
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
CC amplify nucleic acid from an isolate belonging to a specific genotype, or

CC as a probe for specific detection/classification of nucleic acid.
CC Polypeptides encoded by the nucleotides in such compositions may be used
CC for immunisation against HCV, for the detection of antibodies directed
CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
CC region of HCV subtype 3a and is taken from a clone designated BR36-20-
CC 165. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 133 AA;

Query Match 100.0%; Score 616; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.3e-62;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQOYD 60
Db 16 ACSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQOYD 75

Qy 61 EMBCSQAAPYIEQAQVIAHQFKGKVLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 118
Db 76 EMBCSQAAPYIEQAQVIAHQFKGKVLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 133

RESULT 2
AAR63289
ID AAR63289 standard; protein; 133 AA.
XX
AC AAR63289;
XX
XX 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX
XX Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping.
XX
XX Hepatitis C virus type 3.
XX
XX WO9425601-A2.
XX
XX 10-NOV-1994.
XX
XX 27-APR-1994; 94WO-EP001323.
XX
XX 27-APR-1993; 93EP-00401099.
PR 05-AUG-1993; 93EP-00402019.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
XX Maertens G, Stuyver L;
XX
XX WPI; 1994-358277/44.
DR N-PSDB; AAQ78041.
XX
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX
XX Claim 11; Page 127; 404pp; English.
XX
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
CC amplify nucleic acid from an isolate belonging to a specific genotype, or
CC as a probe for specific detection/classification of nucleic acid.
CC Polypeptides encoded by the nucleotides in such compositions may be used
CC for immunisation against HCV, for the detection of antibodies directed
CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4

CC region of HCV subtype 3a and is taken from a clone designated BR36-20-
CC 166. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 133 AA;

Query Match 98.7%; Score 608; DB 2; Length 133;
Best Local Similarity 99.2%; Pred. No. 1.1e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQOYD 60
Db 16 ACSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQOYD 75

Qy 61 EMBCSQAAPYIEQAQVIAHQFKGKVLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 118
Db 76 EMBCSQAAPYIEQAQVIAHQFKGKVLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 133

RESULT 3
AAR63290
ID AAR63290 standard; protein; 133 AA.
XX
AC AAR63290;
XX
XX 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX
XX Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping.
XX
XX Hepatitis C virus type 3.
XX
XX WO9425601-A2.
XX
XX 10-NOV-1994.
XX
XX 27-APR-1994; 94WO-EP001323.
XX
XX 27-APR-1993; 93EP-00401099.
PR 05-AUG-1993; 93EP-00402019.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
XX Maertens G, Stuyver L;
XX
XX WPI; 1994-358277/44.
DR N-PSDB; AAQ78042.
XX
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX
XX Claim 11; Page 128-129; 404pp; English.
XX
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
CC amplify nucleic acid from an isolate belonging to a specific genotype, or
CC as a probe for specific detection/classification of nucleic acid.
CC Polypeptides encoded by the nucleotides in such compositions may be used
CC for immunisation against HCV, for the detection of antibodies directed
CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
CC region of HCV subtype 3a and is taken from a clone designated BR36-20-
CC 165. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 133 AA;

Query Match 98.7%; Score 608; DB 2; Length 133;
 Best Local Similarity 99.2%; Pred. No. 1.1e-61;
 Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 60
 DB 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 75

OY 61 EMECSQAAPYIEQAQVIAHQFKGVLGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
 DB 76 EMECSQAAPYIEQAQVIAHQFKGVLGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 133

RESULT 4
 AAR63390
 ID AAR63390 standard; protein; 209 AA.
 XX AC AAR63390;
 XX 25-MAR-2003 (revised)
 DT 18-AUG-1995 (first entry)
 XX HCV polypeptide sequence.
 XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping.
 XX Hepatitis C virus.
 OS
 XX WO9425601-A2.
 XX 10-NOV-1994.
 XX 27-APR-1994; 94WO-EP001323.
 XX 27-APR-1993; 93EP-00401099.
 PR 05-AUG-1993; 93EP-00402019.
 XX (INNO-) INNOGENETICS NV SA.
 PA Maertens G, Stuyver L;
 PI WPI; 1994-358277/44.
 DR N-PSDB; AAQ78125.
 XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates.
 XX Disclosure; Page 274-275; 404pp; English.
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or
 CC as a probe for specific detection/classification of nucleic acid.
 CC Polypeptides encoded by the nucleotides in such compositions may be used
 CC for immunisation against HCV, for the detection of antibodies directed
 CC against HCV and for serotyping. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX SQ Sequence 209 AA;

Query Match 98.7%; Score 608; DB 2; Length 209;
 Best Local Similarity 99.2%; Pred. No. 1.9e-61;
 Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 60
 DB 92 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 151

OY 61 EMECSQAAPYIEQAQVIAHQFKGVLGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
 DB 152 EMECSQAAPYIEQAQVIAHQFKGVLGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 209

RESULT 5
 AAR37937
 ID AAR37937 standard; protein; 128 AA.
 XX AC AAR37937;
 XX 25-MAR-2003 (revised)
 DT 23-SEP-1993 (first entry)
 XX HCV NS4 region consensus sequence.
 DE Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
 KW NS-5; phylogeny; differentiation; NS-3; core region; type 3.
 XX Synthetic.
 OS
 XX WO9310239-A2.
 PN 27-MAY-1993.
 PD 20-NOV-1992; 92WO-GB002143.
 XX 21-NOV-1991; 91GB-00024696.
 PR 24-JUN-1992; 92GB-00013362.
 XX (COMM-) COMMON SERVICES AGENCY.
 PA Simmonds P, Chan S, Yap PL;
 PI WPI; 1993-182554/22.
 DR N-PSDB; AAQ43111.
 XX DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
 PT diagnosing and treating HCV infection, screening blood samples and
 PT identifying different HCV types.
 XX Disclosure; Fig 9b; 120pp; English.
 XX The sequences given in AAR37932-37 show amino acids 1638-1765 of the NS4
 CC region of hepatitis C virus-3 (HCV-3) samples from 5 blood donors and a
 CC consensus sequence. Analysis of this and other regions of the HCV genome
 CC revealed the existence of three distinct groups of HCV. Analysis of the
 CC region encompassing -255 to -62 of the 5' non coding region (see AAQ43058
 CC -73) showed a difference of 9-14% in the nucleotide sequences between the
 CC three groups. Two of the groups identified were similar to those of HCV
 CC variants termed type 1 and 2, whilst the third appeared to represent a
 CC novel type of virus. Comparison of the NS3 region (see AAR37927-30)
 CC showed a high degree of sequence diversity with type 3 being
 CC phylogenetically different to type 1 and 2. The same degree
 CC differentiation was noted in the NS-5 (see AAR37923-26) and core region
 CC between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct
 CC PN field.)
 XX SQ Sequence 128 AA;

Query Match 97.9%; Score 603; DB 2; Length 128;
 Best Local Similarity 98.3%; Pred. No. 3.8e-61;
 Matches 116; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 60
 DB 10 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 69

OY 61 EMECSQAAPYIEQAQVIAHQFKGVLGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118

Db	70	EMECSSQAAPYIEQAQVIAHOFKEKVLGLLORATQQQAVIEPIVATNWKLEAFWHKH	127
RESULT 6			
AAR37936			
ID	AAR37936	standard; protein; 127 AA.	
XX	AC		
XX	AC	AAR37936;	
XX	DT	25-MAR-2003 (revised)	
DT	23-SEP-1993	(first entry)	
XX			
XX	HCV	NS4 region from donor T1787.	
XX			
KW	Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;		
KW	NS-5; phylogeny; differentiation; NS-3; core region; type 3.		
XX	Hepatitis C virus.		
XX	WO9310239-A2.		
XX	27-MAY-1993.		
XX	20-NOV-1992;	92WO-GB002143.	
XX	21-NOV-1991;	91GB-00024696.	
PR	24-JUN-1992;	92GB-00013362.	
XX	(COMM-) COMMON SERVICES AGENCY.		
XX	Simmonds P, Chan S, Yap PL;		
DR	WPI; 1993-182554/22.		
DR	N-PSDB; AAQ43110.		
XX			
PT	DNA encoding antigenic peptide(s) of new types of hepatitis C virus - f		
PT	diagnosing and treating HCV infection, screening blood samples and		
PT	identifying different HCV types.		
XX	Disclosure; Fig 9b; 120pp; English.		
XX			
CC	The sequences given in AAR37932-37 show amino acids 1638-1765 of the NS3		
CC	region of hepatitis C virus-3 (HCV-3) samples from 5 blood donors and a		
CC	consensus sequence. Analysis of this and other regions of the HCV genome		
CC	revealed the existence of three distinct groups of HCV. Analysis of the		
CC	region encompassing -255 to -62 of the 5' non coding region (see AAQ430		
CC	-75) showed a difference of 9-14% in the nucleotide sequences between t		
CC	three groups. Two of the groups identified were similar to those of HCV		
CC	variants termed type 1 and 2, whilst the third appeared to represent a		
CC	novel type of virus. Comparison of the NS3 region (see AAR37927-30)		
CC	showed a high degree of sequence diversity with type 3 being		
CC	phylogenetically different to type 1 and 2. The same degree		
CC	differentiation was noted in the NS-5 (see AAR37923-26) and core region		
CC	between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct		
CC	PN field.)		
XX			
SQ	Sequence 127 AA;		
	Query Match	97.6%;	Score 601; DB 2; Length 127;
	Best Local Similarity	97.5%;	Pred. No. 6.4e-61;
	Matches 115; Conservative	1; Mismatches	2; Indels 0; Gaps
Qy	1	ACMSADLEVTTSWLLGGVLAALAAAYCLSGVCVWVGHIELGKGPAIVDPKVELVQOYD	6
Db	9	ACMSADLEVTTSWLLGGVLAALAAAYCLSGVCVWVGHIELGKGPAIVDPKVELVQOYD	6
Qy	61	EMECSSQAAPYIEQAQVIAHOFKEKVLGLLORATQQQAVIEPIVATNWKLEAFWHKH	118
Db	69	EMECSSQAAPYIEQAQVIAHOFKEKVLGLLORATQQQAVIEPIVATNWKLEAFWHKH	136
RESULT 7			

DT 25-MAR-2003 (revised)
DT 23-SEP-1993 (first entry)
XX
DE HCV NS4 region from donor T0040.
XX
KW Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
KW NS-5; phylogeny; differentiation; NS-3; core region; type 3.
XX
OS Hepatitis C virus.
XX
PN WO9310239-A2.
XX
PD 27-MAY-1993.
XX
XX 20-NOV-1992; 92WO-GB002143.
PF
XX 21-NOV-1991; 91GB-00024696.
PR
PR 24-JUN-1992; 92GB-00013362.
XX
PA (COMM-) COMMON SERVICES AGENCY.
XX
FI Simmonds P, Chan S, Yap PL;
XX
XX WPI; 1993-182554/22.
DR
DR N-PSDB; AAQ43106.
XX
XX DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
PT diagnosing and treating HCV infection, screening blood samples and
PT identifying different HCV types.
XX
PS Disclosure; Fig 9b; 120pp; English.
XX
XX The sequences given in AAR37932-37 show amino acids 1638-1765 of the NS4
CC region of hepatitis C virus-3 (HCV-3) samples from 5 blood donors and a
CC consensus sequence. Analysis of this and other regions of the HCV genome
CC revealed the existence of three distinct groups of HCV. Analysis of the
CC region encompassing -255 to -62 of the 5' non coding region (see AAQ43058
CC -75) showed a difference of 9-14% in the nucleotide sequences between the
CC three groups. Two of the groups identified were similar to those of HCV
CC variants termed type 1 and 2, whilst the third appeared to represent a
CC novel type of virus. Comparison of the NS3 region (see AAR37927-30)
CC showed a high degree of sequence diversity with type 3 being
CC phylogenetically different to type 1 and 2. The same degree
CC differentiation was noted in the NS-5 (see AAR37923-26) and core region
CC between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 128 AA;
Query Match 97.2%; Score 599; DB 2; Length 128;
Best Local Similarity 97.5%; Pred. No. 1.1e-60;
Matches 115; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYD 60
Db 10 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYD 69
QY 61 EMECSQAAPVIEQAQVIAHOFKGVGLGLLQRAQQQAQVIEPIVTTNWKLEAFWHKH 118
Db 70 EMECSQAAPVIEQAQVIAHOFKGVGLGLLQRAQQQAQVIEPIVATNWOKLETFWHKH 127
RESULT 9
ADX40820
ID ADX40820 standard; protein; 3021 AA.
XX
AC ADX40820;
XX
XX 21-APR-2005 (first entry)
DT
DE HCV polymerase protein #43.
XX
KW Immune stimulation; polymerase; enzyme.

XX Hepatitis C virus.
XX WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
DR
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
SQ Sequence 3021 AA;
Query Match 96.9%; Score 597; DB 9; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.2e-58;
Matches 114; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYD 60
Db 1553 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYD 1712
QY 61 EMECSQAAPVIEQAQVIAHOFKGVGLGLLQRAQQQAQVIEPIVTTNWKLEAFWHKH 118
Db 1713 EMECSQAAPVIEQAQVIAHOFKGVGLGLLQRAQQQAQVIEPIVTTNWKLEAFWHKH 1770
RESULT 10
ADX40823
ID ADX40823 standard; protein; 3021 AA.
XX
AC ADX40823;
XX
XX 21-APR-2005 (first entry)
DT
DE HCV polymerase protein #46.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX Disclosure; Page 388-440; 459pp; English.
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX Sequence 3021 AA;
SQ
Query Match 96.9%; Score 597; DB 9; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.2e-58;
Matches 114; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
Db 1653 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 1712
QY 61 EMECSQAAPYIEQAQVIAHQFKVGLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 118
Db 1713 EMECSQAAPYIEQAQVIAHQFKVGLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 1770
SQ
Query Match 96.9%; Score 597; DB 9; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.2e-58;
Matches 114; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
Db 1653 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 1712
QY 61 EMECSQAAPYIEQAQVIAHQFKVGLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 118
Db 1713 EMECSQAAPYIEQAQVIAHQFKVGLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 1770
RESULT 11
ADX40824
ID ADX40824 standard; protein; 3021 AA.
XX
AC ADX40824;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #47.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX Disclosure; Page 388-440; 459pp; English.
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX Sequence 3021 AA;
SQ

CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX Sequence 3021 AA;
SQ
Query Match 96.9%; Score 597; DB 9; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.2e-58;
Matches 114; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
Db 1653 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 1712
QY 61 EMECSQAAPYIEQAQVIAHQFKVGLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 118
Db 1713 EMECSQAAPYIEQAQVIAHQFKVGLGLLQRAQQQAVIEPIVTTNWQKLEAFWHKH 1770
RESULT 12
AAR63287
ID AAR63287 standard; protein; 133 AA.
XX
AC AAR63287;
XX
DT 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX
DE Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
XX
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping.
XX
OS Hepatitis C virus type 3.
XX
PN WO9425601-A2.
XX
PD 10-NOV-1994.
XX
PF 27-APR-1994; 94WO-EP001323.
XX
PR 27-APR-1993; 93EP-00401099.
PR 05-AUG-1993; 93EP-00402019.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI Maertens G, Stuyver L;
XX
DR WPI; 1994-358277/44.
DR N-PSDB; AAQ78039.
XX
PT New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX
PS Claim 11; Page 123-124; 404pp; English.
XX
CC Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
CC amplify nucleic acid from an isolate belonging to a specific genotype, or
CC as a probe for specific detection/classification of nucleic acid.
CC Polypeptides encoded by the nucleotides in such compositions may be used
CC for immunisation against HCV, for the detection of antibodies directed
CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
CC region of HCV subtype 3a and is taken from a clone designated HD10-1-3.
CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 133 AA;
 Query Match 95.8%; Score 590; DB 2; Length 133;
 Best Local Similarity 93.2%; Pred. No. 1.3e-59;
 Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
 Db 16 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 75

Qy 61 EMESCSQAAPYIEQAQVIAHQFKGVLGLLQRAQQQAVIEPIVTTNNQKLEAFWVHKH 118
 Db 76 EMESCSQAAPYIEQAQVIAHQFKGVLGLLQRAQQQAVIEPIVTTNNQKLEAFWVHKH 133

RESULT 13
 AAR63286
 ID AAR63286 standard; protein; 133 AA.
 XX AC AAR63286;
 XX DT 25-MAR-2003 (revised)
 XX DT 01-AUG-1995 (first entry)
 XX DE Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
 XX KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 XX KW classification; immunisation; prophylaxis; serotyping.
 XX OS Hepatitis C virus type 3.
 XX FN WO9425601-A2.
 XX PD 10-NOV-1994.
 XX PF 27-APR-1994; 94WO-EP001323.
 XX PR 27-APR-1993; 93EP-00401099.
 XX PR 05-AUG-1993; 93EP-00402019.
 XX PA (INNO-) INNOGENETICS NV SA.
 XX PI Maertens G, Stuyver L;
 XX WPI; 1994-358277/44.
 XX N-PSDB; AAQ78038.

XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates.

XX Claim 11; Page 121-122; 40pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR3 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or
 CC as a probe for specific detection/classification of nucleic acid.
 CC Polypeptides encoded by the nucleotides in such compositions may be used
 CC for immunisation against HCV, for the detection of antibodies directed
 CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
 CC region of HCV subtype 3a and is taken from a clone designated HD10-1-25.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 133 AA;
 Query Match 95.8%; Score 590; DB 2; Length 133;

Best Local Similarity 93.2%; Pred. No. 1.3e-59;
 Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
 Db 16 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 75

Qy 61 EMESCSQAAPYIEQAQVIAHQFKGVLGLLQRAQQQAVIEPIVTTNNQKLEAFWVHKH 118
 Db 76 EMESCSQAAPYIEQAQVIAHQFKGVLGLLQRAQQQAVIEPIVTTNNQKLEAFWVHKH 133

RESULT 14
 ADX40813
 ID ADX40813 standard; protein; 3021 AA.
 XX AC ADX40813;
 XX DT 21-APR-2005 (first entry)
 XX DE HCV polymerase protein #36.
 XX KW Immune stimulation; polymerase; enzyme.
 XX OS Hepatitis C virus.
 XX FN WO2005012502-A2.
 XX PD 10-FEB-2005.
 XX PF 29-MAR-2004; 2004WO-US009510.
 XX PR 28-MAR-2003; 2003US-0458026P.
 XX PA (EPIM-) EPIMUNE INC.
 XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
 XX WPI; 2005-132661/14.
 XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
 PT response comprises identifying variants of a peptide epitope 8-11 amino
 PT acids in length comprising primary anchor residues of the same HLA class
 PT I binding motif.
 XX PS Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
 CC epitope which induces an HLA class I CTL response against variants of the
 CC peptide epitope, comprising identifying, from a particular antigen of an
 CC infectious agent, variants of a peptide epitope comprising primary anchor
 CC residues of the same HLA class I binding motif. The method is useful for
 CC identifying a candidate peptide epitope, which induces an HLA class I CTL
 CC response against variants of the peptide epitope. This sequence
 CC represents an HCV polymerase protein used in the scope of the invention.

XX SQ Sequence 3021 AA;
 Query Match 94.0%; Score 579; DB 9; Length 3021;
 Best Local Similarity 93.2%; Pred. No. 1.4e-56;
 Matches 110; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
 Db 1653 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 1712

Qy 61 EMESCSQAAPYIEQAQVIAHQFKGVLGLLQRAQQQAVIEPIVTTNNQKLEAFWVHKH 118
 Db 1713 EMESCSQAAPYIEQAQVIAHQFKGVLGLLQRAQQQAVIEPIVTTNNQKLEAFWVHKH 1770

RESULT 15
 AAR37934

```

ID  AAR37934 standard; protein; 117 AA.
XX
AC  AAR37934;
XX
DT  25-MAR-2003 (revised)
DT  23-SEP-1993 (first entry)
XX
DE  HCV NS4 region from donor T0036.
XX
XX  Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
KW  NS-5; phylogeny; differentiation; NS-3; core region; type 3.
XX
OS  Hepatitis C virus.
XX
PN  W09310239-A2.
XX
PD  27-MAY-1993.
XX
XX  20-NOV-1992; 92MO-GB002143.
XX
XX  21-NOV-1991; 91GB-00024596.
XX  24-JUN-1992; 92GB-00013362.
XX
XX  (COMM-) COMMON SERVICES AGENCY.
XX
XX  Simmonds P, Chan S, Yap PL;
XX
DR  MPI; 1993-182554/22.
DR  N-PSDB; AAQ43108.
XX
XX  DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
PT  diagnosing and treating HCV infection, screening blood samples and
PT  identifying different HCV types.
XX
PS  Disclosure; Fig 9b; 120pp; English.
XX
XX  The sequences given in AAR37932-37 show amino acids 1638-1765 of the NS4
CC  region of hepatitis C virus-3 (HCV-3) samples from 5 blood donors and a
CC  consensus sequence. Analysis of this and other regions of the HCV genome
CC  revealed the existence of three distinct groups of HCV. Analysis of the
CC  region encompassing -255 to -62 of the 5' non coding region (see AAQ43058
CC  -75) showed a difference of 9-14% in the nucleotide sequences between the
CC  three groups. Two of the groups identified were similar to those of HCV
CC  variants termed type 1 and 2, whilst the third appeared to represent a
CC  novel type of virus. Comparison of the NS3 region (see AAR37927-30)
CC  showed a high degree of sequence diversity with type 3 being
CC  phylogenetically different to type 1 and 2. The same degree
CC  differentiation was noted in the NS-5 (see AAR37923-26) and core region
CC  between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct
CC  PN field.)
XX
SQ  Sequence 117 AA;

Query Match      88.5%; Score 545; DB 2; Length 117;
Best Local Similarity 97.2%; Pred. No. 1.6e-54;
Matches 106; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy  1  ACMSADLEVTITSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYD 60
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  9  ACMSADLEVTITSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYD 68
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy  61  EMECSQAAPYIEQAQVIAHQPKVGLGLQRAATQQAQVIEPIVTTNWQ 109
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  69  EMECSQAAPYIEQAQVIAHQPKVGLGLQRAATQQAQVIEPIVATNWQ 117
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:07:29 ; Search time 45 Seconds
(without alignments)
216.794 Million cell updates/sec

Title: US-09-638-693a-36_COPY_16_133
Perfect score: 616
Sequence: 1 ACMSADLEVTSTWVLLGGV.....VIEPIVTTNWQKLEAFWHKH 118

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	133	2	US-09-878-281A-36
2	608	98.7	133	2	US-09-878-281A-38
3	608	98.7	133	2	US-09-878-281A-40
4	608	98.7	209	2	US-09-878-281A-223
5	590	95.8	133	2	US-09-878-281A-32
6	590	95.8	133	2	US-09-878-281A-34
7	542	88.0	128	1	US-08-244-116B-17
8	467	75.8	3012	2	US-08-811-566-2
9	467	75.8	3012	2	US-09-034-756-2
10	462	75.0	313	1	US-08-483-695-45
11	462	75.0	313	1	US-07-965-285-45
12	462	75.0	313	1	US-08-487-231-45
13	462	75.0	313	2	US-09-201-912-45
14	462	75.0	360	2	US-08-850-328-4
15	462	75.0	382	2	US-08-444-818-68
16	462	75.0	460	2	US-08-444-818-20
17	462	75.0	592	2	US-08-867-611-47
18	462	75.0	592	2	US-09-690-359-47
19	462	75.0	594	2	US-08-867-611-48
20	462	75.0	594	2	US-09-690-359-48
21	462	75.0	597	2	US-08-867-611-16
22	462	75.0	597	2	US-09-690-359-16
23	462	75.0	597	4	PCT-US92-06965A-21
24	462	75.0	599	2	US-08-867-611-18
25	462	75.0	599	2	US-09-690-359-18
26	462	75.0	599	4	PCT-US92-06965A-23
27	462	75.0	613	2	US-08-867-611-49

28	462	75.0	613	2	US-09-690-359-49	Sequence 49, Appl
29	462	75.0	739	2	US-08-444-818-148	Sequence 148, Appl
30	462	75.0	859	2	US-08-444-818-30	Sequence 30, Appl
31	462	75.0	971	2	US-08-867-611-52	Sequence 52, Appl
32	462	75.0	971	2	US-09-690-359-52	Sequence 52, Appl
33	462	75.0	973	2	US-08-867-611-53	Sequence 53, Appl
34	462	75.0	973	2	US-09-690-359-53	Sequence 53, Appl
35	462	75.0	992	2	US-08-867-611-54	Sequence 54, Appl
36	462	75.0	992	2	US-09-690-359-54	Sequence 54, Appl
37	462	75.0	1021	1	US-07-910-760-12	Sequence 12, Appl
38	462	75.0	1021	1	US-08-440-519-12	Sequence 12, Appl
39	462	75.0	1021	2	US-08-440-519-12	Sequence 12, Appl
40	462	75.0	2261	2	US-08-444-818-66	Sequence 66, Appl
41	462	75.0	2436	2	US-08-444-818-75	Sequence 75, Appl
42	462	75.0	2772	2	US-08-444-818-89	Sequence 89, Appl
43	462	75.0	2894	1	US-08-466-975A-23	Sequence 23, Appl
44	462	75.0	2894	1	US-08-391-671A-23	Sequence 23, Appl
45	462	75.0	2894	2	US-08-467-902A-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1

US-09-878-281A-36
; Sequence 36, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:

; APPLICANT: Innogenetics N.V.

; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35

; CURRENT APPLICATION NUMBER: US/09/878,281A

; CURRENT FILING DATE: 2001-06-12

; NUMBER OF SEQ ID NOS: 284

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 36

; LENGTH: 133

; TYPE: PRT

; ORGANISM: hepatitis C virus

US-09-878-281A-36

Query Match 100.0%; Score 616; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.2e-65;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIGHIELGKPKATVPDKVLYQQYD	60
Db	16	ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIGHIELGKPKATVPDKVLYQQYD	75
Qy	61	EMBECSQAAPYIEQAQVIAHQFKGKVLGLLQRAIQQAQVIEPIVTTNWQKLEAFWHKH	118
Db	76	EMBECSQAAPYIEQAQVIAHQFKGKVLGLLQRAIQQAQVIEPIVTTNWQKLEAFWHKH	133

RESULT 2

US-09-878-281A-38
; Sequence 38, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:

; APPLICANT: Innogenetics N.V.

; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35

; CURRENT APPLICATION NUMBER: US/09/878,281A

; CURRENT FILING DATE: 2001-06-12

; NUMBER OF SEQ ID NOS: 284

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 38

; LENGTH: 133

; TYPE: PRT

; ORGANISM: hepatitis C virus

US-09-878-281A-38

```
Query Match      98.7%; Score 608; DB 2; Length 133;
Best Local Similarity 99.2%; Pred. No. 2.8e-64;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
    |||||
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 75
    |||||
Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 118
    |||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 133
    |||||

RESULT 3
US-09-878-281A-40
; Sequence 40, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 133
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-40

Query Match      98.7%; Score 608; DB 2; Length 133;
Best Local Similarity 99.2%; Pred. No. 2.8e-64;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
    |||||
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 75
    |||||
Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 118
    |||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 133
    |||||

RESULT 4
US-09-878-281A-223
; Sequence 223, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
; LENGTH: 209
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-223

Query Match      98.7%; Score 608; DB 2; Length 209;
Best Local Similarity 99.2%; Pred. No. 5.2e-64;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
    |||||
Db 92 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 151
    |||||
```

```
Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 118
    |||||
Db 152 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 209
    |||||

RESULT 5
US-09-878-281A-32
; Sequence 32, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 133
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-32

Query Match      95.8%; Score 590; DB 2; Length 133;
Best Local Similarity 93.2%; Pred. No. 3.9e-62;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
    |||||
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 75
    |||||
Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 118
    |||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 133
    |||||

RESULT 6
US-09-878-281A-34
; Sequence 34, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 133
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-34

Query Match      95.8%; Score 590; DB 2; Length 133;
Best Local Similarity 93.2%; Pred. No. 3.9e-62;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
    |||||
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 75
    |||||
Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 118
    |||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLLQORATQQAQVIEPIVTTNWQKLEAFWHKH 133
    |||||

RESULT 7
US-08-244-116B-17
; Sequence 17, Application US/08244116B
; Patent No. 5763159
```



```
;
; GENERAL INFORMATION:
; APPLICANT: Simmonds, Peter
; APPLICANT: Chan, Shiu-Wan
; APPLICANT: Yap, Peng L.
; TITLE OF INVENTION: Hepatitis-C Virus Testing
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5763159th Carolina
; COUNTRY: United States
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244,116B
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02143
; FILING DATE: 20-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 1749-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 704-377-1561
; TELEFAX: 704-334-2014
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: yes
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis-C virus
;
; US-08-244-116B-17
;
; Query Match 88.0%; Score 542; DB 1; Length 128;
; Best Local Similarity 97.2%; Pred. No. 1.8e-56;
; Matches 106; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
;
; Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGGKPAIVDPKDEVLYQQYD 60
; Db 10 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGGKPAIVDPKDEVLYQQYD 69
;
; Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLGLLQORATQQQAVIEPIVTTNQ 109
; Db 70 EMECSQAAPYIEQAQVIAHQFKVGLGLLQORATQQQAVIEPIVATNQ 118
;
; RESULT 8
; US-08-811-566-2
; Sequence 2, Application US/08811566
; Patent No. 6127116
; GENERAL INFORMATION:
; APPLICANT: Rice, Charles et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
;
; Query Match 88.0%; Score 467; DB 2; Length 3012;
; Best Local Similarity 75.2%; Pred. No. 9.5e-46;
; Matches 88; Conservative 9; Mismatches 20; Indels 0; Gaps 0;
;
; Qy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGGKPAIVDPKDEVLYQQYDE 61
; Db 1648 CMSADLEVTSTWLLGGVLAALAAAYCLSTGCVVIGHIELGGKPAIIPDREVLVQEFDE 1707
;
; Qy 62 MEECSQAAPYIEQAQVIAHQFKVGLGLLQORATQQQAVIEPIVTTNQKLEAFMHHK 118
; Db 1708 MEECSQAAPYIEQAQVIAHQFKVGLGLLQORATQQQAVIEPIVATNQKLEAFMHHK 1764
;
; RESULT 9
; US-09-034-756-2
; Sequence 2, Application US/09034756
; Patent No. 6392028
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,756
; FILING DATE: 04-May-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
;
```

```
;
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,566
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1113-1-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3012 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
;
; US-08-811-566-2
;
; Query Match 75.8%; Score 467; DB 2; Length 3012;
; Best Local Similarity 75.2%; Pred. No. 9.5e-46;
; Matches 88; Conservative 9; Mismatches 20; Indels 0; Gaps 0;
;
; Qy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGGKPAIVDPKDEVLYQQYDE 61
; Db 1648 CMSADLEVTSTWLLGGVLAALAAAYCLSTGCVVIGHIELGGKPAIIPDREVLVQEFDE 1707
;
; Qy 62 MEECSQAAPYIEQAQVIAHQFKVGLGLLQORATQQQAVIEPIVTTNQKLEAFMHHK 118
; Db 1708 MEECSQAAPYIEQAQVIAHQFKVGLGLLQORATQQQAVIEPIVATNQKLEAFMHHK 1764
;
; RESULT 9
; US-09-034-756-2
; Sequence 2, Application US/09034756
; Patent No. 6392028
; GENERAL INFORMATION:
; APPLICANT: RICE, CHARLES et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWELL & HAFERKAMP, L.C.
; STREET: 7733 FORSYTH BLVD., SUITE 1400
; CITY: ST. LOUIS
; STATE: MO
; COUNTRY: USA
; ZIP: 63105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,756
; FILING DATE: 04-May-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: HOLLAND, DONALD R.
; REGISTRATION NUMBER: 35,197
; REFERENCE/DOCKET NUMBER: 6029-4831
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-727-5188
;
```

```
TELEFAX: 314-727-6092
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 3012 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-034-756-2

Query Match      75.8%; Score 467; DB 2; Length 3012;
Best Local Similarity 75.2%; Pred. No. 9.5e-46;
Matches 88; Conservative 9; Mismatches 20; Indels 0; Gaps 0;

Qy      2 CMSADLEVTSTWVLGGVLAALAAAYCLSGCVVIVGHIELGGKPAIVPDKVLYQQYDE 61
Db      1648 CMSADLEVTSTWVLGGVLAALAAAYCLSGCVVIVGHIELGGKPAIVPDKVLYQQYDE 1707

Qy      62 MEECSQAAPYIEQAVIAHOFKGVLLGLLQRTAQQAIVIEPIVTTNNQKLEAFWHKH 118
Db      1708 MEECSQHLPIEQGMMLAEQFKQKALGLLQRTASRQAEVITPAVQTNQKLEVFPAKH 1764

RESULT 10
US-08-483-695-45
; Sequence 45, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 313 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; Query Match      75.0%; Score 462; DB 1; Length 313;
; Best Local Similarity 73.5%; Pred. No. 1.8e-46;
; Matches 86; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
```

```
US-08-483-695-45
; Sequence 45, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 313 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; Query Match      75.0%; Score 462; DB 1; Length 313;
; Best Local Similarity 73.5%; Pred. No. 1.8e-46;
; Matches 86; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
```

RESULT 12
US-08-487-231-45
; Sequence 45, Application US/08487231
; Patent No. 5919454
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,231
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-487-231-45

Query Match 75.0%; Score 462; DB 1; Length 313;
Best Local Similarity 73.5%; Pred. No. 1.8e-46;
Matches 86; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
Qy 2 CMSADLEVTSTWVLGGVLAALAAAYCISVGCVVIVGHIELGKPAIVDPKVELVQOYDE 61
Db 194 CMSADLEVTSTWVLGGVLAALAAAYCISVGCVVIVGHIELGKPAIIPDREVLVRSFDE 253
Qy 62 MEECSQAAPYIEQAVIAHQFKGVLGLLQRTAQOAVIEPIVTNNWQLEAFWHKH 118
Db 254 MEECSQHLPIYIEQGMWLAEQFKQKALGLLQRTASQAQVIAPAVETNNWQKLETFWAKH 310

RESULT 13
US-09-201-912-45
; Sequence 45, Application US/09201912
; Patent No. 6210962
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadorf, Dina

; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESS: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,285
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-201-912-45

Query Match 75.0%; Score 462; DB 2; Length 313;
Best Local Similarity 73.5%; Pred. No. 1.8e-46;
Matches 86; Conservative 11; Mismatches 20; Indels 0; Gaps 0;
Qy 2 CMSADLEVTSTWVLGGVLAALAAAYCISVGCVVIVGHIELGKPAIVDPKVELVQOYDE 61
Db 194 CMSADLEVTSTWVLGGVLAALAAAYCISVGCVVIVGHIELGKPAIIPDREVLVRSFDE 253
Qy 62 MEECSQAAPYIEQAVIAHQFKGVLGLLQRTAQOAVIEPIVTNNWQLEAFWHKH 118
Db 254 MEECSQHLPIYIEQGMWLAEQFKQKALGLLQRTASQAQVIAPAVETNNWQKLETFWAKH 310

RESULT 14
US-08-850-328-4
; Sequence 4, Application US/08850328
; Patent No. 6379886
; GENERAL INFORMATION:
; APPLICANT: TAKAHAMA, Y.
; APPLICANT: SHIRAI, J.
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS
; TITLE OF INVENTION: C VIRUS INFECTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ FOR WINDOWS Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,328
FILING DATE: 02-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mays, Thomas D
REGISTRATION NUMBER: 34,524
REFERENCE/DOCKET NUMBER: 32273-20004.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 360 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-850-328-4

Query Match 75.0%; Score 462; DB 2; Length 360;
Best Local Similarity 73.5%; Pred. No. 2.2e-46;
Matches 86; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

Qy 2 CMSADLEVTTSWVLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVPDKEVLYQOYDE 61
Db 8 CMSADLEVTTSWVLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVPDKEVLYQOYDE 67

Qy 62 MEECSQAAPYIEQAOVIAHOFKGVLGGLLQATQQOAVIEPIVTTNNQKLEAFWHKH 118
Db 68 MEECSQHLPLYIEQGMMLAEQFKQKALGLLQATASRQAEVIAPAVQTNQKLETFWAKH 124

RESULT 15
US-08-444-818-68
Sequence 68, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876

TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-68

Query Match 75.0%; Score 462; DB 2; Length 382;
Best Local Similarity 73.5%; Pred. No. 2.4e-46;
Matches 86; Conservative 11; Mismatches 20; Indels 0; Gaps 0;

Qy 2 CMSADLEVTTSWVLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVPDKEVLYQOYDE 61
Db 94 CMSADLEVTTSWVLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVPDKEVLYQOYDE 153

Qy 62 MEECSQAAPYIEQAOVIAHOFKGVLGGLLQATQQOAVIEPIVTTNNQKLEAFWHKH 118
Db 154 MEECSQHLPLYIEQGMMLAEQFKQKALGLLQATASRQAEVIAPAVQTNQKLETFWAKH 210

Search completed: January 27, 2006, 19:15:13
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:13:35 ; Search time 116 Seconds
(without alignments)
425.033 Million cell updates/sec

Title: US-09-638-693a-36_COPY_16_133
Perfect score: 616
Sequence: 1 ACMGADLEVTSTWLLGGV.....VIEPIVTTNMOKLEAFWHKH 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:
1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgm2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
3: /cgm2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
4: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
5: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
6: /cgm2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	100.0	133	3	US-09-899-046-36
2	616	100.0	133	3	US-09-878-281-36
3	616	100.0	133	3	US-09-873-224-36
4	608	98.7	133	3	US-09-899-046-38
5	608	98.7	133	3	US-09-899-046-40
6	608	98.7	133	3	US-09-878-281-38
7	608	98.7	133	3	US-09-878-281-40
8	608	98.7	133	3	US-09-873-224-38
9	608	98.7	133	3	US-09-873-224-40
10	608	98.7	209	3	US-09-899-046-223
11	608	98.7	209	3	US-09-878-281-223
12	608	98.7	209	3	US-09-873-224-223
13	590	95.8	133	3	US-09-899-046-32
14	590	95.8	133	3	US-09-899-046-34
15	590	95.8	133	3	US-09-878-281-32
16	590	95.8	133	3	US-09-878-281-34
17	590	95.8	133	3	US-09-873-224-32
18	590	95.8	133	3	US-09-873-224-34
19	542	88.0	128	4	US-10-396-964-17
20	467	75.8	3011	3	US-09-742-659-4
21	467	75.8	3011	3	US-09-891-894-3
22	467	75.8	3011	4	US-10-184-150-3
23	467	75.8	3011	4	US-10-328-997-3
24	467	75.8	3012	3	US-09-238-076-2
25	467	75.8	3012	3	US-09-995-937-2
26	467	75.8	3012	3	US-09-917-563-2
27	463	75.2	3011	4	US-10-296-734-06

28	462	75.0	360	6	US-11-126-662-4	Sequence 4, Appli
29	462	75.0	1892	5	US-10-612-884-6	Sequence 6, Appli
30	462	75.0	2894	3	US-09-941-611-23	Sequence 23, Appli
31	462	75.0	2894	4	US-10-044-995-23	Sequence 23, Appli
32	462	75.0	2894	5	US-10-822-871-23	Sequence 23, Appli
33	462	75.0	3011	3	US-09-916-359-2	Sequence 2, Appli
34	462	75.0	3011	3	US-09-238-076-20	Sequence 20, Appli
35	462	75.0	3011	3	US-09-952-572-9	Sequence 9, Appli
36	462	75.0	3011	3	US-09-929-955-1	Sequence 1, Appli
37	462	75.0	3011	3	US-09-747-419-20	Sequence 20, Appli
38	462	75.0	3011	3	US-09-995-937-20	Sequence 20, Appli
39	462	75.0	3011	3	US-09-917-563-20	Sequence 20, Appli
40	462	75.0	3011	4	US-10-104-966-1	Sequence 1, Appli
41	462	75.0	3011	4	US-10-259-275-20	Sequence 20, Appli
42	462	75.0	3011	4	US-10-232-643-6	Sequence 6, Appli
43	462	75.0	3011	4	US-10-189-359-14	Sequence 14, Appli
44	462	75.0	3011	4	US-10-719-619-1	Sequence 1, Appli
45	462	75.0	3011	4	US-10-445-724-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-899-046-36
; Sequence 36, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-899-046-36

Query Match 100.0%; Score 616; DB 3; Length 133;
Best Local Similarity 100.0%; Pred. No. 6.7e-62;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ACMGADLEVTSTWLLGGVLAALAAVCLSGCVVIVGHIELGKGPATVPDKVLYQQVD	60
Db	16	ACMGADLEVTSTWLLGGVLAALAAVCLSGCVVIVGHIELGKGPATVPDKVLYQQVD	75
Qy	61	EMECSSQAAPYIEQAQVIAHQFKGKVLGLLQATQQQAIVIEPIVTTNMOKLEAFWHKH	118
Db	76	EMECSSQAAPYIEQAQVIAHQFKGKVLGLLQATQQQAIVIEPIVTTNMOKLEAFWHKH	133

RESULT 2
US-09-878-281-36
; Sequence 36, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/878,281
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/362,455
FILING DATE:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 133 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-878-281-36

Query Match      100.0%; Score 616; DB 3; Length 133;
Best Local Similarity 100.0%; Pred. No. 6.7e-62;
Matches 116; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAMSADLEVTSTWVLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKVELYQYD 60
    |||||||
Db 16 AAMSADLEVTSTWVLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKVELYQYD 75
    |||||||

Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLGLQRATQQQAVIEPIVTTNWQKLEAFWHKH 118
    |||||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLGLQRATQQQAVIEPIVTTNWQKLEAFWHKH 133
    |||||||

RESULT 3
US-09-873-224-36
; Sequence 36, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Innogenetics sa.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-09-873-224-36

Query Match      100.0%; Score 616; DB 3; Length 133;
```

```
Best Local Similarity 100.0%; Pred. No. 6.7e-62;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAMSADLEVTSTWVLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKVELYQYD 60
    |||||||
Db 16 AAMSADLEVTSTWVLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKVELYQYD 75
    |||||||

Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLGLQRATQQQAVIEPIVTTNWQKLEAFWHKH 118
    |||||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLGLQRATQQQAVIEPIVTTNWQKLEAFWHKH 133
    |||||||

RESULT 4
US-09-899-046-38
; Sequence 38, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-899-046-38

Query Match      98.7%; Score 608; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 5.4e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 AAMSADLEVTSTWVLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKVELYQYD 60
    |||||||
Db 16 AAMSADLEVTSTWVLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKVELYQYD 75
    |||||||

Qy 61 EMECSQAAPYIEQAQVIAHQFKVGLGLQRATQQQAVIEPIVTTNWQKLEAFWHKH 118
    |||||||
Db 76 EMECSQAAPYIEQAQVIAHQFKVGLGLQRATQQQAVIEPIVTTNWQKLEAFWHKH 133
    |||||||

RESULT 5
US-09-899-046-40
; Sequence 40, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
```

```
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 133 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-899-046-40

Query Match      98.7%; Score 608; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 5.4e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVPDKEVLYQQYD 60
DB 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVPDKEVLYQQYD 75
QY 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQQRATQQQAVIEPIVTTNNQKLEAFWHKH 118
DB 76 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQQRATQQQAVIEPIVTTNNQKLEAFWHKH 133

RESULT 6
US-09-878-281-38
; Sequence 38, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/878,281
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 133 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-878-281-38

Query Match      98.7%; Score 608; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 5.4e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVPDKEVLYQQYD 60
DB 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVPDKEVLYQQYD 75
QY 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQQRATQQQAVIEPIVTTNNQKLEAFWHKH 118
DB 76 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQQRATQQQAVIEPIVTTNNQKLEAFWHKH 133

RESULT 7
US-09-878-281-40
; Sequence 40, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/878,281
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 133 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-09-878-281-40

Query Match      98.7%; Score 608; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 5.4e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVPDKEVLYQQYD 60
DB 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVPDKEVLYQQYD 75
QY 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQQRATQQQAVIEPIVTTNNQKLEAFWHKH 118
DB 76 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQQRATQQQAVIEPIVTTNNQKLEAFWHKH 133

RESULT 8
US-09-873-224-38
; Sequence 38, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
; TITLE OF INVENTION: genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Innogenetics SA.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 133 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-873-224-38

Query Match      98.7%; Score 608; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 5.4e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy 1 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 16 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 75
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 118
Db 76 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 133

RESULT 9
US-09-873-224-40
; Sequence 40, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
; genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Innogenetics sa.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-873-224-40

Query Match 98.7%; Score 608; DB 3; Length 133;
Best Local Similarity 99.2%; Pred. No. 5.4e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 16 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 75
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 118
Db 76 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 133

RESULT 10
US-09-899-046-223
; Sequence 223, Application US/09899046
; Publication No. US2003008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 223:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-899-046-223

Query Match 98.7%; Score 608; DB 3; Length 209;
Best Local Similarity 99.2%; Pred. No. 9.7e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 92 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 151
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 118
Db 152 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 209

RESULT 11
US-09-878-281-223
; Sequence 223, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; genotypes for diagnosis, prophylaxis and therapy.
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/878,281
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 223:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-878-281-223

Query Match 98.7%; Score 608; DB 3; Length 209;
Best Local Similarity 99.2%; Pred. No. 9.7e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 92 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 151
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 118
Db 152 EMECSQAAPYIEQAQVIAHQFKGVGLLQQRATQQQAVIEPIVTTNWQKLEAFWPKH 209

RESULT 12

US-09-873-224-223
; Sequence 223, Application US/09873224
; Publication No. US20030064360A1
; GENERAL INFORMATION:
; APPLICANT: <Unknown>
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; CORRESPONDENCE ADDRESS:
; STREET: Industriepark Zwijnaarde 7, box 4
; CITY: Ghent
; COUNTRY: Belgium
; ZIP: B-9052
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/873,224
; FILING DATE: 05-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: ImmoGenetics sa.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 00 32 9 241 07 11
; TELEFAX: 00 32 9 241 07 99
; INFORMATION FOR SEQ ID NO: 223:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 209 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 223:
US-09-873-224-223
Query Match 98.7%; Score 608; DB 3; Length 209;
Best Local Similarity 99.2%; Pred. No. 9.7e-61;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 60
Db ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 151
Qy 61 EMESCSQAAPYIEQAQVIAHQFKGKVLGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 118
Db 152 EMESCSQAAPYIEQAQVIAHQFKKVLGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 209
RESULT 13
US-09-899-046-32
; Sequence 32, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-899-046-32
Query Match 95.8%; Score 590; DB 3; Length 133;
Best Local Similarity 93.2%; Pred. No. 6.1e-59;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 60
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 75
Qy 61 EMESCSQAAPYIEQAQVIAHQFKGKVLGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 118
Db 76 EMESCSQAAPYIEQAQVIAHQFKKILGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 133
RESULT 14
US-09-899-046-34
; Sequence 34, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-899-046-34
Query Match 95.8%; Score 590; DB 3; Length 133;
Best Local Similarity 93.2%; Pred. No. 6.1e-59;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 60
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 75
Qy 61 EMESCSQAAPYIEQAQVIAHQFKGKVLGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 118
Db 76 EMESCSQAAPYIEQAQVIAHQFKKILGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 133
RESULT 15
US-09-878-281-32
; Sequence 32, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-899-046-32
Query Match 95.8%; Score 590; DB 3; Length 133;
Best Local Similarity 93.2%; Pred. No. 6.1e-59;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 60
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 75
Qy 61 EMESCSQAAPYIEQAQVIAHQFKGKVLGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 118
Db 76 EMESCSQAAPYIEQAQVIAHQFKKILGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 133
RESULT 14
US-09-899-046-34
; Sequence 34, Application US/09899046
; Publication No. US20030008274A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/899,046
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-899-046-34
Query Match 95.8%; Score 590; DB 3; Length 133;
Best Local Similarity 93.2%; Pred. No. 6.1e-59;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 60
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQQYD 75
Qy 61 EMESCSQAAPYIEQAQVIAHQFKGKVLGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 118
Db 76 EMESCSQAAPYIEQAQVIAHQFKKILGLLQRLQATQQQAIVIEPIVTTNNWOKLEAFWHKH 133
RESULT 15
US-09-878-281-32
; Sequence 32, Application US/09878281
; Publication No. US20030032005A1
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: New sequences of hepatitis C virus
; NUMBER OF SEQUENCES: 270
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/878,281
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/362,455
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 133 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-878-281-32

Query Match      95.8%; Score 590; DB 3; Length 133;
Best Local Similarity 93.2%; Pred. No. 6.le-59;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy      1 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQOYD 60
Db      16 ACSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHIELGKPAIVDPKVELYQOYD 75

Qy      61 EMECSQAAPYIEQAQVIAHQFKGVLGLLQRAVIEPIVTTWQKLEAFWKKH 118
Db      76 EMECSQAAPYIEQAQVIAHQFKGVLGLLQRAVIEPIVTTWQKLEAFWKKH 133

Search completed: January 27, 2006, 19:25:10
Job time : 117 secs

```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	462	75.0	3011	6	US-10-985-205-3	Sequence 3, Appli
2	423	68.7	1985	7	US-11-119-330-2	Sequence 2, Appli
3	422	68.5	2280	7	US-11-022-562-211	Sequence 211, App
4	227	36.9	54	7	US-11-137-220-2	Sequence 2, Appli
5	76	12.3	23	7	US-11-135-414-31	Sequence 31, Appl
6	64	10.4	831	6	US-10-467-657-4486	Sequence 4486, Ap
7	61	9.9	257	7	US-11-159-597-6	Sequence 6, Appli
8	61	9.9	257	7	US-11-219-359-1	Sequence 1, Appli
9	61	9.9	275	7	US-11-159-597-8	Sequence 8, Appli
10	61	9.9	277	7	US-11-159-597-4	Sequence 4, Appli
11	61	9.9	279	7	US-11-159-597-10	Sequence 10, Appl
12	61	9.9	401	6	US-10-467-657-1860	Sequence 1860, Ap
13	61	9.9	943	6	US-10-467-657-5508	Sequence 5508, Ap
14	60.5	9.8	759	7	US-11-149-003-22	Sequence 22, Appl
15	60.5	9.8	1057	7	US-11-149-003-6	Sequence 6, Appli
16	60.5	9.8	1192	7	US-11-149-003-18	Sequence 18, Appl
17	60.5	9.8	1207	7	US-11-149-003-20	Sequence 20, Appl
18	60.5	9.8	1251	7	US-11-149-003-16	Sequence 16, Appl
19	60.5	9.8	1342	7	US-11-149-003-24	Sequence 24, Appl
20	60.5	9.8	1477	7	US-11-149-003-8	Sequence 8, Appli
21	60.5	9.8	1512	7	US-11-149-003-10	Sequence 10, Appl
22	60.5	9.8	1535	7	US-11-149-003-14	Sequence 14, Appl
23	60.5	9.8	1570	7	US-11-149-003-12	Sequence 12, Appl
24	60.5	9.8	1593	7	US-11-149-003-4	Sequence 4, Appli
25	60.5	9.8	1628	7	US-11-149-003-2	Sequence 2, Appli

```
; CURRENT FILING DATE: 2005-04-29
; PRIOR APPLICATION NUMBER: 60/567,270
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 60/568,590
; PRIOR FILING DATE: 2004-05-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon
US-11-119-330-2

Query Match 68.7%; Score 423; DB 7; Length 1985;
Best Local Similarity 66.9%; Pred. No. 6.8e-39;
Matches 79; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWLVGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
Db 622 ACSADLEVTSTWLVGGVLAALAAAYCLTGTGTVIVGRILSGKPAIIPDREVLRYEFD 681
Qy 61 EMECSQAAPVIEQAQVIAHOFKGVGLGLLQRTAQOQAVIEPIVTTNWOKLEAFWKH 118
Db 682 EMECASHLPVIEQGMQLAEQFKQKAGLLQTAIKQAEAAAPVVESKRWTLAEAFWAKH 739

RESULT 3
US-11-022-562-211
; Sequence 211, Application US/11022562
; Publication No. US20050249742A1
; GENERAL INFORMATION:
; APPLICANT: Ruprecht, Ruth M.
; APPLICANT: Shisong, Jiang
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATING
; TITLE OF INVENTION: A CYTOTOXIC T LYMPHOCYTE IMMUNE RESPONSE
; FILE REFERENCE: DFN-043CN
; CURRENT APPLICATION NUMBER: US/11/022,562
; CURRENT FILING DATE: 2004-12-22
; PRIOR APPLICATION NUMBER: PCT/US03/20322
; PRIOR FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 60/392718
; PRIOR FILING DATE: 2002-06-27
; NUMBER OF SEQ ID NOS: 340
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 211
; LENGTH: 2280
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-11-022-562-211

Query Match 68.5%; Score 422; DB 7; Length 2280;
Best Local Similarity 66.9%; Pred. No. 1e-38;
Matches 79; Conservative 14; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWLVGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYD 60
Db 1647 ACSADLEVTSTWLVGGVLAALAAAYCLTGTGTVIVGRILSGKPAVVPDREVLRYEFD 1706
Qy 61 EMECSQAAPVIEQAQVIAHOFKGVGLGLLQRTAQOQAVIEPIVTTNWOKLEAFWKH 118
Db 1707 EMECASHLPVIEQGMQLAEQFKQKAGLLQTAIKQAEAAAPVVESKRWTLAEAFWAKH 1764

RESULT 4
US-11-137-220-2
; Sequence 2, Application US/11137220
; Publication No. US2005027192A1
; GENERAL INFORMATION:
; APPLICANT: Saliberg, Matti
; APPLICANT: TRIPEP AB
; TITLE OF INVENTION: HEPATITIS C VIRUS NONSTRUCTURAL PROTEIN
```

```
; TITLE OF INVENTION: 4A (NS4A) IS AN ENHANCER ELEMENT
; FILE REFERENCE: TRIPEP.052CI
; CURRENT APPLICATION NUMBER: US/11/137,220
; CURRENT FILING DATE: 2005-05-24
; PRIOR APPLICATION NUMBER: PCT/IB2003/006488
; PRIOR FILING DATE: 2003-11-25
; PRIOR APPLICATION NUMBER: 60/430,009
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-11-137-220-2

Query Match 36.9%; Score 227; DB 7; Length 54;
Best Local Similarity 77.8%; Pred. No. 4.9e-19;
Matches 42; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Qy 12 STWLVGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQOYDEMEEC 65
Db 1 STWLVGGVLAALAAAYCLTGTGTVIVGRILSGKPAIIPDREVLRYEFDMEEC 54

RESULT 5
US-11-135-414-31
; Sequence 31, Application US/11135414
; Publication No. US20050266498A1
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, SATORU
; APPLICANT: MIWA, KIYOSHI
; APPLICANT: ETO, YUZURU
; TITLE OF INVENTION: METHOD FOR SCREENING BIOMOLECULE ACTIVITY REGULATOR
; FILE REFERENCE: 213701USOPCT
; CURRENT APPLICATION NUMBER: US/11/135,414
; CURRENT FILING DATE: 2005-05-24
; PRIOR APPLICATION NUMBER: US/09/936,179
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: JP 99/11-63110
; PRIOR FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 23
; TYPE: PRT
; ORGANISM: ARTIFICIAL SEQUENCE
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-11-135-414-31

Query Match 12.3%; Score 76; DB 7; Length 23;
Best Local Similarity 65.2%; Pred. No. 0.011;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 29 LSVGCVVIVGHIELGKPAIVDP 51
Db 1 LTTGTVIVGRILSGKPAVVPD 23

RESULT 6
US-10-467-657-4486
; Sequence 4486, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
```

```
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 486
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-4486

Query Match          10.4%; Score 64; DB 6; Length 831;
Best Local Similarity 25.3%; Pred. No. 24;
Matches 25; Conservative 17; Mismatches 39; Indels 18; Gaps 4;

Qy 11 TSTWLLGGV-----LAALAAAYCLSVGCVWIVGHIELGGKPAIVDPKVELVYQQ--- 58
Db 468 TALMITVALLVITPCALSLATPTALAAGSTGLAREG-ILIGGKQAI-----ETLSQTTDI 522

Qy 59 -YDEMECSQAAPYIEQAQVIAHQFQKGVGLGLLQRAQQ 96
Db 523 IFDKTGTLTQGNPAVRRIELGSMTEAQVLAQAQSLQQ 561

RESULT 7
US-11-159-597-6
; Sequence 6, Application US/11159597
; Publication No. US20050255559A1
; GENERAL INFORMATION:
; APPLICANT: Uebele, Victor N.
; APPLICANT: Swanson, Richard J.
; APPLICANT: Liu, Yuan
; APPLICANT: Lagrutta, Armando
; TITLE OF INVENTION: NOVEL HUMAN CALCIUM SENSITIVE POTASSIUM
; TITLE OF INVENTION: CHANNEL
; FILE REFERENCE: 20499P
; CURRENT APPLICATION NUMBER: US/11/159,597
; CURRENT FILING DATE: 2005-06-23
; PRIOR APPLICATION NUMBER: US/10/031,691
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/19585
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/144,764
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-159-597-6

Query Match          9.9%; Score 61; DB 7; Length 257;
Best Local Similarity 30.0%; Pred. No. 12;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

Qy 16 LGGVLAALAAAYCLSVGCVWIVGHIELGGKPAIVDPKVELVYQQYDEM--ECSQAAPYIE 73
Db 199 LGGALI-----VGMVRLTQHLSL-----LCEKYSTVVRDEVGKVPYIE 238

Qy 74 QAQVIAHQFQKGVGLLQRA 93
Db 239 Q-----HQFK---LCIMERS 250

RESULT 8
US-11-159-359-1
; Sequence 1, Application US/11219359
; Publication No. US20060004187A1
; GENERAL INFORMATION:
; APPLICANT: Jeglia, Timothy James
; APPLICANT: Wickenden, Alan
; APPLICANT: Liu, Yi
US-11-159-359-1

Query Match          9.9%; Score 61; DB 7; Length 275;
Best Local Similarity 30.0%; Pred. No. 13;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

Qy 16 LGGVLAALAAAYCLSVGCVWIVGHIELGGKPAIVDPKVELVYQQYDEM--ECSQAAPYIE 73
Db 217 LGGALI-----VGMVRLTQHLSL-----LCEKYSTVVRDEVGKVPYIE 256

Qy 74 QAQVIAHQFQKGVGLLQRA 93
```

```
; APPLICANT: ICAGEN, Inc.
; TITLE OF INVENTION: BK Beta Subunits of Slo Family Potassium Channels
; FILE REFERENCE: 018512-002030US
; CURRENT APPLICATION NUMBER: US/11/219,359
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US/09/914,053
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/121,224
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: US 60/163,367
; PRIOR FILING DATE: 1999-11-03
; PRIOR APPLICATION NUMBER: WO PCT/US00/04441
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: BK beta 2
US-11-219-359-1

Query Match          9.9%; Score 61; DB 7; Length 257;
Best Local Similarity 30.0%; Pred. No. 12;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

Qy 16 LGGVLAALAAAYCLSVGCVWIVGHIELGGKPAIVDPKVELVYQQYDEM--ECSQAAPYIE 73
Db 199 LGGALI-----VGMVRLTQHLSL-----LCEKYSTVVRDEVGKVPYIE 238

Qy 74 QAQVIAHQFQKGVGLLQRA 93
Db 239 Q-----HQFK---LCIMERS 250

RESULT 9
US-11-159-597-8
; Sequence 8, Application US/11159597
; Publication No. US20050255559A1
; GENERAL INFORMATION:
; APPLICANT: Uebele, Victor N.
; APPLICANT: Swanson, Richard J.
; APPLICANT: Liu, Yuan
; APPLICANT: Lagrutta, Armando
; TITLE OF INVENTION: NOVEL HUMAN CALCIUM SENSITIVE POTASSIUM
; TITLE OF INVENTION: CHANNEL
; FILE REFERENCE: 20499P
; CURRENT APPLICATION NUMBER: US/11/159,597
; CURRENT FILING DATE: 2005-06-23
; PRIOR APPLICATION NUMBER: US/10/031,691
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/US00/19585
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/144,764
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 275
; TYPE: PRT
; ORGANISM: Amino Acid
US-11-159-597-8

Query Match          9.9%; Score 61; DB 7; Length 275;
Best Local Similarity 30.0%; Pred. No. 13;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

Qy 16 LGGVLAALAAAYCLSVGCVWIVGHIELGGKPAIVDPKVELVYQQYDEM--ECSQAAPYIE 73
Db 217 LGGALI-----VGMVRLTQHLSL-----LCEKYSTVVRDEVGKVPYIE 256

Qy 74 QAQVIAHQFQKGVGLLQRA 93
```

```
Db      257 Q-----HQFK---LCIMRS 268
      ||||| |:::|
Query Match      9.9%; Score 61; DB 7; Length 279;
Best Local Similarity 30.0%; Pred. No. 13;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

RESULT 10
US-11-159-597-4
; Sequence 4, Application US/11159597
; Publication No. US20050255559A1
; GENERAL INFORMATION:
; APPLICANT: Uebele, Victor N.
; APPLICANT: Swanson, Richard J.
; APPLICANT: Liu, Yuan
; APPLICANT: Lagrutta, Armando
; TITLE OF INVENTION: NOVEL HUMAN CALCIUM SENSITIVE POTASSIUM
; FILE REFERENCE: 20499P
; CURRENT FILING DATE: 2005-06-23
; PRIOR FILING DATE: 2005-06-23
; PRIOR FILING DATE: US/10/031,691
; PRIOR FILING DATE: 2002-01-22
; PRIOR FILING DATE: PCT/US00/19585
; PRIOR FILING DATE: 2000-07-18
; PRIOR FILING DATE: 2000-07-18
; PRIOR FILING DATE: 60/144,764
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-11-159-597-4

Query Match      9.9%; Score 61; DB 7; Length 277;
Best Local Similarity 30.0%; Pred. No. 13;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

QY      16 LLGGVLAALAAAYCLSVGVVGHIELGGKPAIVPDKEVLYQOYDEM--EBCSQAAPYIE 73
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      219 LLGGALI-----VGMVRLTQHLSL-----LCEKYSTVVRDEVGKVPYIE 259
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      74 QAOVIAHQPKVGLLQRA 93
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      259 Q-----HQFK---LCIMRS 270
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

RESULT 11
US-11-159-597-10
; Sequence 10, Application US/11159597
; Publication No. US20050255559A1
; GENERAL INFORMATION:
; APPLICANT: Uebele, Victor N.
; APPLICANT: Swanson, Richard J.
; APPLICANT: Liu, Yuan
; APPLICANT: Lagrutta, Armando
; TITLE OF INVENTION: NOVEL HUMAN CALCIUM SENSITIVE POTASSIUM
; FILE REFERENCE: 20499P
; CURRENT FILING DATE: 2005-06-23
; PRIOR FILING DATE: 2005-06-23
; PRIOR FILING DATE: US/10/031,691
; PRIOR FILING DATE: 2002-01-22
; PRIOR FILING DATE: PCT/US00/19585
; PRIOR FILING DATE: 2000-07-18
; PRIOR FILING DATE: 2000-07-18
; PRIOR FILING DATE: 60/144,764
; PRIOR FILING DATE: 1999-07-20
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Amino Acid
US-11-159-597-10

Query Match      9.9%; Score 61; DB 7; Length 279;
Best Local Similarity 30.0%; Pred. No. 13;
Matches 24; Conservative 10; Mismatches 16; Indels 30; Gaps 5;

QY      16 LLGGVLAALAAAYCLSVGVVGHIELGGKPAIVPDKEVLYQOYDEM--EBCSQAAPYIE 73
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      219 LLGGALI-----VGMVRLTQHLSL-----LCEKYSTVVRDEVGKVPYIE 259
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      74 QAOVIAHQPKVGLLQRA 93
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      259 Q-----HQFK---LCIMRS 270
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

RESULT 12
US-10-467-657-1860
; Sequence 1860, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1860
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1860

Query Match      9.9%; Score 61; DB 6; Length 401;
Best Local Similarity 28.4%; Pred. No. 21;
Matches 27; Conservative 15; Mismatches 27; Indels 26; Gaps 6;

QY      35 VIVGHIELGGKPAIVPDKEVLYQOYD--EMECSSQAAPYIEQAQVIAHQPKVGLLQRA 92
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      137 VIVVH--LQMPAEMDGMALAKEHDLWVIEDCAQAH-----GATYKSGVSGIGH 185
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

QY      93 ATQ-----QAVIEP-----IVTTN-----WQKLEAF 114
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|
Db      186 VGAWSPQDKIITGGEGGNVTTNDKTLWEKWAY 220
      ||||| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::| |:::|

RESULT 13
US-10-467-657-5508
; Sequence 5508, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 5508
; LENGTH: 943
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-5508
```

```
Query Match          9.9%; Score 61; DB 6; Length 943;
Best Local Similarity 35.6%; Pred. No. 61;
Matches 21; Conservative 8; Mismatches 24; Indels 6; Gaps 3;

QY 37 VGHIELGKPAIVDPKEVLYQQYDEMECSQAAPYIEQAQVIAHQPKGK--VLGLLQRA 93
DB 414 LGHI---GRPAANDPEAFLEGGAEE--AEALPRPPVVTVMGVHDHGKTSLLDYIRRA 468

RESULT 14
US-11-149-003-22
; Sequence 22, Application US/11149003
; Publication No. US20060014277A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Novel Human Kielin-like Proteins and Polynucleotides Encoding the
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/11/149,003
; PRIOR FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US/10/189,971
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 759
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-149-003-22

Query Match          9.8%; Score 60.5; DB 7; Length 759;
Best Local Similarity 23.6%; Pred. No. 53;
Matches 34; Conservative 11; Mismatches 64; Indels 35; Gaps 5;

QY 2 CWSADLEY-----TTSTWVLGGVLAALAAAYCLSVGCVVIVGH----- 39
DB 436 CHSGDFS VHVNTNDRGRSGVAWTQEVAVLLGD-----MAVRLLDQGA VTDGHPVALPFLQ 491
QY 40 -----IELGGKPAIV---PDKEVLYQQYDEMECSQAAPYIEQAQVIAHQPKGKVLGLLQ 91
DB 492 EPLLVEVLRGHTVILHAQFGLQVLWDGOSQV-EVSVPGSYQGRTCGLCGNFNGFAQDDLIQ 550
QY 92 RATQQQAVIEPIVTTNWOKLEAFW 115
DB 551 GPEGLLPSEAAFGNSQVSEGLW 574

RESULT 15
US-11-149-003-6
; Sequence 6, Application US/11149003
; Publication No. US20060014277A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Novel Human Kielin-like Proteins and Polynucleotides Encoding the
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0360-USA
; CURRENT APPLICATION NUMBER: US/11/149,003
; PRIOR FILING DATE: 2005-06-09
; PRIOR APPLICATION NUMBER: US/10/189,971
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/302,949
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/315,634
; PRIOR FILING DATE: 2001-08-29
```

```
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1057
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-149-003-6
```

```
Query Match          9.8%; Score 60.5; DB 7; Length 1057;
Best Local Similarity 23.6%; Pred. No. 81;
Matches 34; Conservative 11; Mismatches 64; Indels 35; Gaps 5;

QY 2 CWSADLEY-----TTSTWVLGGVLAALAAAYCLSVGCVVIVGH----- 39
DB 734 CHSGDFS VHVNTNDRGRSGVAWTQEVAVLLGD-----MAVRLLDQGA VTDGHPVALPFLQ 789
QY 40 -----IELGGKPAIV---PDKEVLYQQYDEMECSQAAPYIEQAQVIAHQPKGKVLGLLQ 91
DB 790 EPLLVEVLRGHTVILHAQFGLQVLWDGOSQV-EVSVPGSYQGRTCGLCGNFNGFAQDDLIQ 848
QY 92 RATQQQAVIEPIVTTNWOKLEAFW 115
DB 849 GPEGLLPSEAAFGNSQVSEGLW 872
```

```
Search completed: January 27, 2006, 19:25:53
Job time : 32 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:05:23 ; Search time 39 Seconds
(without alignments)
291.117 Million cell updates/sec

Title: US-09-638-693a-36_COPY_16_133

Perfect score: 616
Sequence: 1 ACMSADLEVTSTWVLGGV.....VIEPIVTNNQKLEAFWKKH 118

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_80.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	608	98.7	209	2 PC1306	genome polyprotein
2	590	95.8	142	2 PC1307	genome polyprotein
3	464	75.3	3011	1 S40770	genome polyprotein
4	462	75.0	3011	1 GNWVC3	genome polyprotein
5	462	75.0	3011	1 GNWVCH	genome polyprotein
6	460	74.7	492	2 PS0326	polyprotein - hepa
7	457	74.2	716	2 J01366	polyprotein - hepa
8	426	69.2	3010	1 S18030	genome polyprotein
9	422	68.5	3010	1 A45573	genome polyprotein
10	421	68.3	3010	1 GNWVCJ	genome polyprotein
11	420	68.2	3010	1 GNWVTC	genome polyprotein
12	418	67.9	3010	1 GNWVTW	genome polyprotein
13	411	66.7	3014	1 JC5620	genome polyprotein
14	363	58.9	3033	1 J01303	genome polyprotein
15	362	58.8	3033	1 GNWJ38	genome polyprotein
16	351	57.0	876	2 PC2219	polyprotein - hepa
17	346	56.2	125	2 SC3529	hypothetical prote
18	132	21.4	41	2 PQ0560	nonstructural prot
19	131	21.3	41	2 PQ0563	nonstructural prot
20	131	21.3	41	2 PQ0564	nonstructural prot
21	131	21.3	41	2 PQ0562	nonstructural prot
22	131	21.3	41	2 PQ0565	nonstructural prot
23	113	18.3	41	2 PQ0561	nonstructural prot
24	83	13.5	1108	2 AF1047	probable membrane
25	78	12.7	343	2 S75435	hypothetical prote
26	77	12.5	1107	2 B91271	probable periplasm
27	77	12.5	1107	2 B86112	probable periplasm
28	77	12.5	1107	2 E65226	hypothetical 123.8
29	74.5	12.1	346	2 B97014	TPR-repeat-contain

ALIGNMENTS

RESULT 1

PC1306

genome polyprotein NS4a epitope containing region (isolate BR36-20) - hepatitis C virus
C:Species: hepatitis C virus
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: PC1306

R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; DeLeys, R.; Maertens, G.

Biochem. Biophys. Res. Commun. 192, 635-641, 1993

A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.
A:Reference number: PC1300; MUID:93249436; PMID:7683463

A:Accession: PC1306

A:Molecule type: mRNA

A:Residues: 1-209 <STU>

A:Cross-references: UNIPROT:Q81594; UNIPARC:UPI0000033846; DDBJ:D14600; NID:9303584; PDB:

A:Experimental source: blood

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: nonstructural protein; polyprotein

Query Match 98.7%; Score 608; DB 2; Length 209;

Best Local Similarity 99.2%; Pred. No. 5.2e-55;

Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAVCLSGVCVIVGHIELGKPAIVDPKEVLYQQYD 60

Db 92 ACMSADLEVTSTWVLGGVLAALAAVCLSGVCVIVGHIELGKPAIVDPKEVLYQQYD 151

Qy 61 EMESCSQAAPVIEQAQVIAHQFKGLGLLQRAQQQAVIEPIVTNNQKLEAFWKKH 118

Db 152 EMESCSQAAPVIEQAQVIAHQFKGLGLLQRAQQQAVIEPIVTNNQKLEAFWKKH 209

RESULT 2

PC1307

genome polyprotein NS4a epitope containing region (isolate HD10-1) - hepatitis C virus (

C:Species: hepatitis C virus

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 31-Dec-2004

C:Accession: PC1307

R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; DeLeys, R.; Maertens, G.

Biochem. Biophys. Res. Commun. 192, 635-641, 1993

A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.

A:Reference number: PC1300; MUID:93249436; PMID:7683463

A:Accession: PC1307

A:Molecule type: mRNA

A:Residues: 1-142 <STU>

A:Cross-references: UNIPROT:Q68870; UNIPARC:UPI0000178536; DDBJ:D14602

A:Experimental source: blood

C:Keywords: polyprotein

Query Match 95.8%; Score 590; DB 2; Length 142;

Best Local Similarity 93.2%; Pred. No. 2.4e-53;

Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

F:116-191/Product: capsid protein C #status predicted <CPC>
F:192-389/Product: major envelope protein M #status predicted <EPM>
F:390-729/Product: major envelope protein E #status predicted <MEE>
F:730-1006/Product: nonstructural protein NS1 #status predicted <NS1>
F:1007-1615/Product: nonstructural protein NS2 #status predicted <NS2>
F:1230-1237/Product: hepatitis C virus #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

Query Match 75.0%; Score 462; DB 1; Length 3011;
Best Local Similarity 74.4%; Pred. No. 1.1e-38;
Matches 87; Conservative 9; Mismatches 21; Indels 0; Gaps 0;

Qy 2 CMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYDE 61
Db 1648 CMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYDE 1707

Qy 62 MEECSQAAPYIEQAOVIAHQFKGVKGLLQORATQQAQVIEPIVTNNWKLQLEAFWKKH 118
Db 1708 MEECSQHLPIYIEQGMMLAEQFKQKALGLLQTRSRQAEVITPAVQTNWQLEAFWAKH 1764

RESULT 6
PS0326
polyprotein - hepatitis C virus (isolate Fla) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C:Accession: PS0326
R:Li, J.S.; Tong, S.P.; Vitvitski, L.; Lepot, D.; Trepo, C.
Gene 105, 167-172, 1991
A:Title: Two French genotypes of hepatitis C virus: homology of the predominant genotype
A:Reference number: PS0326; MUID:92039028; PMID:1719820
A:Accession: PS0326
A:Molecule type: genomic RNA
A:Residues: 1-492 <L1D>
A:Cross-references: UNIPROT:Q91PE5; UNIPROT:Q36579; UNIPROT:Q36610; UNIPROT:Q03463; UNIPARC:UPI0000178532; GB:M60220
A:Note: this sequence corresponds to nonstructural protein NS3 region
A:Note: translation of the nucleotide sequence is not complete
C:Keywords: polyprotein

Query Match 74.7%; Score 460; DB 2; Length 492;
Best Local Similarity 73.5%; Pred. No. 2.2e-39;
Matches 86; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

Qy 2 CMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYDE 61
Db 199 CMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYDE 258

Qy 62 MEECSQAAPYIEQAOVIAHQFKGVKGLLQORATQQAQVIEPIVTNNWKLQLEAFWKKH 118
Db 259 MEECSQHLPIYIEQGMMLAEQFKQKALGLLQTRSRQAEVITPAVQTNWQLEAFWAKH 315

RESULT 7
JQ1366
polyprotein - hepatitis C virus (French isolate) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C:Accession: JQ1366
R:Kremsdorff, D.; Porchon, C.; Kim, J.P.; Reyes, G.R.; Brechot, C.
J. Gen. Virol. 72, 2557-2561, 1991
A:Title: Partial nucleotide sequence analysis of a French hepatitis C virus: implication
A:Reference number: JQ1366; MUID:92013977; PMID:1655961
A:Accession: JQ1366
A:Molecule type: genomic RNA
A:Residues: 1-716 <KRE>
A:Cross-references: UNIPROT:Q9PX22; UNIPARC:UPI0000178531

C:Keywords: glycoprotein; polyprotein
F:84,90,97,115,143,199,223,243,290,312/Binding site: carbohydrate (asn) (covalent) #status
Query Match 74.2%; Score 457; DB 2; Length 716;
Best Local Similarity 72.6%; Pred. No. 6.9e-39;
Matches 85; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

Qy 2 CMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYDE 61
Db 597 CMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYDE 656

Qy 62 MEECSQAAPYIEQAOVIAHQFKGVKGLLQORATQQAQVIEPIVTNNWKLQLEAFWKKH 118
Db 657 MEECSQHLPIYIEQGMMLAEQFKQKALGLLQTRSRQAEVITPAVQTNWQLEAFWAKH 713

RESULT 8
SI8030
genome polyprotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5)
C:Species: hepatitis C virus
A:Variety: isolate JK1
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: SI8030; S33570; A48332; SI8029
R:Honda, M.; Kaneko, S.; Magashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient
A:Reference number: SI8028
A:Accession: SI8030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:Q68949; UNIPARC:UPI00000F2A81; EMBL:X61596; NID:G59478; PIDN:
A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Uoudra, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated f
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA
A:Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HOW>
A:Cross-references: UNIPARC:UPI00001749FF; EMBL:X61591
A:Note: this sequence is inconsistent with the nucleotide translation
A:Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as Trp, and TTC for residue 771 as Ser
A:Note: sequence extracted from NCBI backbone (NCBI:121747, NCBI:121748)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (Asr)

Query Match 69.2%; Score 426; DB 1; Length 3010;
Best Local Similarity 67.8%; Pred. No. 5.4e-35;
Matches 80; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYD 60
Db 1647 ACSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIIPDKREVLYQOYD 1706

Qy 61 EMECSQAAPYIEQAOVIAHQFKGVKGLLQORATQQAQVIEPIVTNNWKLQLEAFWKKH 118
Db 1707 EMECSQHLPIYIEQGMMLAEQFKQKALGLLQTRSRQAEVITPAVQTNWQLEAFWAKH 1764

Db 1707 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1764

RESULT 12
GNWVTW
genome polypotein - hepatitis C virus (strain Taiwan)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain Taiwan)
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A40244
R:Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the
A:Reference number: A40244; MUID:92230206; PMID:1314449
A:Accession: A40244
A:Molecule type: genomic RNA
A:Residues: 1-3010 <CH>
A:Cross-references: UNIPROT:P29846; UNIPARC:UPI0000131E2B; GB:M84754
C:Superfamily: hepatitis C virus genome polypotein
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: major envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <BPM>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypotein
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196, 209, 233, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 207

Query Match 67.9%; Score 418; DB 1; Length 3010;
Best Local Similarity 68.7%; Pred. No. 3.6e-34;
Matches 79; Conservative 12; Mismatches 24; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHIELGKGPATVPDKEVLYQQYD 60
Db 1647 ACMSADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHIELGKGPATVPDKEVLYQQYD 1706

Qy 61 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1761
Db 1707 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1761

RESULT 13
JC5620
genome polypotein - hepatitis C virus (isolate EUH1480)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (isolate EUH1480)
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: JC5620
R:Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A:Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A:Reference number: JC5620; MUID:97366593; PMID:9223423
A:Accession: JC5620
A:Molecule type: mRNA
A:Residues: 1-3014 <CHA>
A:Cross-references: UNIPROT:O39928; UNIPARC:UPI0000174A01; GB:Y13184
A:Experimental source: genotype 5a, which predominates in South Africa
A:Note: the translation of the nucleotide sequence is not complete in this paper
C:Superfamily: hepatitis C virus genome polypotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: major envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <BPM>
F:384-408/Region: hypervariable #status predicted

F:390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F:731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F:1008-1616/Product: hepatitis C virus genome polypotein
F:1231-1238/Region: nucleotide-binding motif A (P-loop)
F:1313-1318/Region: nucleotide-binding motif B
F:1317-1320/Region: DEXH motif
F:1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
F:1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
F:2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F:2210-2249/Region: interferon sensitivity determining #status predicted

Query Match 66.7%; Score 411; DB 1; Length 3014;
Best Local Similarity 65.3%; Pred. No. 1.9e-33;
Matches 77; Conservative 16; Mismatches 25; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHIELGKGPATVPDKEVLYQQYD 60
Db 1648 ACMSADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHIELGKGPATVPDKEVLYQQYD 1707

Qy 61 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1765
Db 1708 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1765

RESULT 14
JQ1303
genome polypotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (isolate HC-J6)
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: JQ1303
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Kural, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.;
J. Gen. Virol. 72, 2697-2704, 1991
A:Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human
A:Reference number: JQ1303; MUID:92044440; PMID:1658196
A:Accession: JQ1303
A:Molecule type: genomic RNA
A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26660; UNIPARC:UPI0000131E25; GB:D00944; NID:9221650; PIDN:1
A:Experimental source: isolate HC-J6 from a Japanese individual
C:Superfamily: hepatitis C virus genome polypotein
C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polypotein; serine proteinase; transme
F:2-115/Product: capsid protein M #status predicted <CPC>
F:116-191/Product: major envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <BPM>
F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F:1011-1619/Product: hepatitis C virus genome polypotein
F:1316-1321/Region: nucleotide-binding motif B
F:1320-1323/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F:196, 209, 234, 305, 325, 417, 423, 430, 448, 477, 534, 542, 558, 578, 627, 649, 1091, 1217, 1259, 2038, 281

Query Match 58.9%; Score 363; DB 1; Length 3033;
Best Local Similarity 56.4%; Pred. No. 1.6e-28;
Matches 66; Conservative 18; Mismatches 33; Indels 0; Gaps 0;

Qy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHIELGKGPATVPDKEVLYQQYDE 61
Db 1652 CMQADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHIELGKGPATVPDKEVLYQQYDE 1711

Qy 62 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1768
Db 1712 EMEBCASHLPYIEQGMQLAEQFKQKALGLLQATATQKAEAAAPVBSKRWALETFPAKH 1768

RESULT 15
GNWVJ8
genome polypotein - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain HC-J8)

protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C/Species: hepatitis C virus
C/Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text_change 09-Jul-2004
C/Accession: A0250; PQ0397; PQ0559
R/Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.;
Virology 188, 331-341, 1992
A/Title: Full-length sequence of a hepatitis C virus genome having poor homology to repc
A/Reference number: A0250; MUID:92230232; PMID:131459
A/Accession: A0250
A/Molecule type: genomic RNA
A/Residues: 1-3033 <OKA>
A/Cross-references: UNIPROT:P26661; UNIPARC:UPI0000131E27; GB:D10988; GB:D01221; NID:G22
R/Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I
J. Gen. Virol. 73, 1131-1141, 1992
A/Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A/Reference number: PQ0393; MUID:92268871; PMID:1316939
A/Accession: PQ0397
A/Molecule type: genomic RNA
A/Residues: 2678-2754 <CHA>
A/Cross-references: UNIPARC:UPI0000174A02; DDBJ:D10134
A/Experimental source: isolate E-bl2
R/Kato, N.; Ootsuyama, Y.; Okoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A/Title: Distribution of plural HCV types in Japan.
A/Reference number: PQ0554; MUID:92068204; PMID:1720309
A/Accession: PQ0559
A/Molecule type: mRNA
A/Residues: 2678-2729 <KAT>
A/Cross-references: UNIPARC:UPI00000F5263; GB:D10562; GB:D90518; NID:G221523; PIDN:BA001
C/Superfamily: hepatitis C virus genome polyprotein
A/Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
P/1-115/Product: capsid protein C #status predicted <CPC>
P/116-191/Product: envelope protein M #status predicted <EPM>
P/192-389/Product: major envelope protein E #status predicted <MEE>
P/390-733/Product: nonstructural protein NS1 #status predicted <NS1>
P/734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
P/1011-1619/Product: hepatitis C virus protein NS3 #status predicted <NS3>
P/1234-1241/Region: nucleotide-binding motif A (P-loop)
P/1316-1321/Region: nucleotide-binding motif B
P/1320-1323/Region: DEXH motif
P/1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
P/1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
P/2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
P/196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,23

Query Match 58.8%; Score 362; DB 1; Length 3033;
Best Local Similarity 53.8%; Pred. No. 2e-28;
Matches 63; Conservative 22; Mismatches 32; Indels 0; Gaps 0;

Qy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKVELYQQYDE 61
Db 1652 CMQADLEIMTSSWVLAGGVLAALAAAYCLATGCTSIIGRLHNDRVVAVDPKELLYEAFDE 1711

Qy 62 MEECSQAAPYIQAVIAHOFKGVLLGQATQQAQVIEPIVTNNWKLKLEAFWHKH 118
Db 1712 MEECSKAALIEEQRMWMLKSKIQGLLQATQQAQVIEPIVTNNWKLKLEAFWHKH 1768

Search completed: January 27, 2006, 19:14:16
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:04:23 ; Search time 161 Seconds
(without alignments)

517.095 Million cell updates/sec

Title: US-09-638-693A-36_COPY_16_133

Perfect score: 616

Sequence: 1 ACMGADLEVTSTWLLGGV.....VIEPIVTNNQKLEAFWHKH 118

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	608	98.7	209	2	Q81594_9HEPC	Q81594 hepatitis c
2	597	96.9	3021	2	O92933_9HEPC	O92933 hepatitis c
3	597	96.9	3021	2	Q68870_9HEPC	Q68870 hepatitis c
4	597	96.9	3021	2	Q81258_9HEPC	Q81258 hepatitis c
5	590	95.8	133	2	Q81595_9HEPC	Q81595 hepatitis c
6	589	95.6	138	2	Q68233_9HEPC	Q68233 hepatitis c
7	579	94.0	3021	2	Q81495_9HEPC	Q81495 hepatitis c
8	577	93.7	138	2	Q68239_9HEPC	Q68239 hepatitis c
9	572	92.9	138	2	Q68241_9HEPC	Q68241 hepatitis c
10	539	87.5	193	2	O56637_9HEPC	O56637 hepatitis c
11	527	85.6	3023	2	Q81487_9HEPC	Q81487 hepatitis c
12	481	78.1	3019	2	Q68801_9HEPC	Q68801 hepatitis c
13	473	76.8	138	2	Q68223_9HEPC	Q68223 hepatitis c
14	470	76.3	658	2	Q68K66_9HEPC	Q68K66 hepatitis c
15	467	75.8	138	2	Q68224_9HEPC	Q68224 hepatitis c
16	467	75.8	659	2	Q68K49_9HEPC	Q68K49 hepatitis c
17	466	75.6	138	2	Q68225_9HEPC	Q68225 hepatitis c
18	466	75.6	659	2	Q68K35_9HEPC	Q68K35 hepatitis c
19	465	75.5	659	2	Q68K39_9HEPC	Q68K39 hepatitis c
20	465	75.5	2908	2	Q61X04_9HEPC	Q61X04 hepatitis c
21	464	75.3	3011	2	Q03463_9HEPC	Q03463 hepatitis c
22	463	75.2	138	2	Q68226_9HEPC	Q68226 hepatitis c
23	463	75.2	659	2	Q68K45_9HEPC	Q68K45 hepatitis c
24	463	75.2	659	2	Q68K43_9HEPC	Q68K43 hepatitis c
25	462	75.0	659	2	Q68K34_9HEPC	Q68K34 hepatitis c
26	462	75.0	659	2	Q68K56_9HEPC	Q68K56 hepatitis c
27	462	75.0	659	2	Q68K41_9HEPC	Q68K41 hepatitis c
28	462	75.0	659	2	Q68K40_9HEPC	Q68K40 hepatitis c
29	462	75.0	2436	2	Q81756_9HEPC	Q81756 hepatitis c
30	462	75.0	3010	1	POLG_HCV1	P26664 h genome po
31	462	75.0	3010	1	POLG_HCVH	P27958 h genome po

RESULT 1									
Q81594_9HEPC	Q81594_9HEPC PRELIMINARY;	PRT;	209 AA.						
AC	Q81594;								
DT	01-NOV-1996 (TrEMBLrel. 01, Created)								
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)								
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)								
DE	Nonstructural protein 4 (Fragment).								
GN	Name=NS4;								
OS	Hepatitis C virus.								
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;								
OC	Hepacivirus.								
OX	NCBI_TaxID=11103;								
RN	[1]								
RP	NUCLEOTIDE SEQUENCE.								
RX	MEDLINE=93249436; PubMed=7683463;								
RA	Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;								
RT	"Analysis of the putative E1 envelope and NS4a epitope region of HCV								
RT	type 3.";								
RL	Biochem. Biophys. Res. Commun. 192:635-641(1993).								
DR	EMBL; D14600; BAA03449.1; -; Genomic_RNA.								
DR	PIR; PCL306; PCL306.								
DR	HSP; P26663; ICUL.								
DR	SMR; Q81594; 1-102.								
DR	InterPro; IPR000745; HCV NS4a.								
DR	InterPro; IPR001490; HCV NS4b.								
DR	Pfam; PF01006; HCV NS4a; 1.								
DR	Pfam; PF01001; HCV NS4b; 1.								
FT	NON_TER 1 1								
FT	NON_TER 209 209								
SQ	SEQUENCE 209 AA; 23408 MW; 76648D9BB1D3CD12 CRC64;								
Query Match 98.7%; Score 608; DB 2; Length 209;									
Best Local Similarity 99.2%; Pred. No. 8.6e-53;									
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
Qy	1 ACMGADLEVTSTWLLGGVLAALAAAYCLSGCVVIVGHIELGKPAIVPDKVLYQQYD 60								
Db	92 ACMGADLEVTSTWLLGGVLAALAAAYCLSGCVVIVGHIELGKPAIVPDKVLYQQYD 151								
Qy	61 EMESCSQAAPYIEQAQVIAHQKGLVGLLQRAQQQAVIEPIVTNNQKLEAFWHKH 118								
Db	152 EMESCSQAAPYIEQAQVIAHQKGLVGLLQRAQQQAVIEPIVTNNQKLEAFWHKH 209								
RESULT 2									
O92933_9HEPC	O92933_9HEPC PRELIMINARY;	PRT;	3021 AA.						
AC	O92933;								
DT	01-NOV-1998 (TrEMBLrel. 08, Created)								
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)								
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)								
DE	Polyprotein.								
OS	Hepatitis C virus.								

```
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP STRAIN-type 3a;
RC NUCLEOTIDE SEQUENCE.
RA Shukla D.D., Chaturvedi S., Cao J.Y., Hoynes P.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF046986; AAC03058.1; -, Genomic_RNA.
DR HSSP: Q8JYS1; 1CWX.
DR SMR: Q92933; 1035-1663, 2431-2996.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_NS5a.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR SMART: SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3021 AA; 328905 MW; D7B6133B330303CD CRC64;
Query Match 96.9%; Score 597; DB 2; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.7e-50;
Matches 114; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 ACMGADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHILGKPKALVPDKVLYQQYD 60
Db 1653 ACMGADLEVTSTWLLGGVLAALAAAYCLSGVGVVGHILGKPKALVPDKVLYQQYD 1712
Qy 61 EMESCSQAAPYIEQAQVIAHQFKGVLLGLLQATQQCAVIEPIVTNNQKLEAFWHKH 118
Db 1713 EMESCSQAAPYIEQAQVIAHQFKGVLLGLLQATQQCAVIEPIVTNNQKLEAFWHKH 1770
RESULT 3
Q68870_9HEPC
ID Q68870_9HEPC PRELIMINARY; PRT; 3021 AA.
AC Q68870;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Genes for core, envelope and NS1 proteins.
OS Hepatitis C virus.
```

```
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Seelig R., Weber P., Seeling H.P., Ledger N., Bottner C., Renz M.;
RL "Hepatitis C virus type V genome isolated from a patient in Germany.";
RT Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93249436; PubMed=7683463;
RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;
RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV type 3.";
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92288871; PubMed=1316939;
RA Chan S., McOmish F., Holmes E., Dow B., Peutherer J., Follett E.,
Yap P., Simmonds P.;
RT "Analysis of a new hepatitis C virus type and its phylogenetic relationship to existing variants.";
RL J. Gen. Virol. 73:1131-1141(1992).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C virus: identification of a new virus type and restrictions on sequence diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL: X76918; CAA54244.1; -, mRNA.
DR PIR: PC1307; PC1307.
DR PIR: PQ0401; PQ0401.
DR PIR: PQ0804; PQ0804.
DR PIR: S41288; S41288.
DR HSSP: Q8JYS1; 1CWX.
DR SMR: Q68870; 1035-1663, 2431-2996.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV_capsid.
DR InterPro: IPR002521; HCV_core.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NS1.
DR InterPro: IPR000745; HCV_NS4a.
DR InterPro: IPR001490; HCV_NS4b.
DR InterPro: IPR002868; HCV_NS5a.
DR InterPro: IPR002166; HCV_NS5a.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV_capsid; 1.
DR Pfam: PF01542; HCV_core; 1.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NS1; 1.
DR Pfam: PF01538; HCV_NS2; 1.
DR Pfam: PF02907; HCV_NS3; 1.
DR Pfam: PF01006; HCV_NS4a; 1.
DR Pfam: PF01506; HCV_NS5a; 1.
DR SMART: SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3021 AA; 328905 MW; D7B6133B330303CD CRC64;
```



```

DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Envelope protein.
FT CHAIN 1 191
FT CHAIN 192 383
FT CHAIN 384 735
FT CHAIN 736 1012
FT CHAIN 1013 1663
FT CHAIN 1664 1717
FT CHAIN 1718 1978
FT CHAIN 1979 2430
FT CHAIN 2431 3021
SQ SEQUENCE 3021 AA; 329096 MW; BF2B499AA55A58CF CRC64;

Query Match 96.9%; Score 597; DB 2; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.7e-50;
Matches 114; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQQYD 60
Db 1653 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQQYD 1712

Qy 61 EMESCSQAAPYIEQAQVIAHQFKGVLLGLLQRAQQQAQVIEPIVTTNNQKLEAFWHKH 118
Db 1713 EMESCSQAAPYIEQAQVIAHQFKGVLLGLLQRAQQQAQVIEPIVATNNQKLEAFWHKH 1770

RESULT 4
Q81258_9HEPC
ID Q81258_9HEPC PRELIMINARY; PRT; 3021 AA.
AC Q81258;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NZL1;
RA Sakamoto M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NZL1;
RA Okamoto H.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92268871; PubMed=8316939;
RA Chan S., McOmish P., Holmes E., Dow B., Peutherer J., Follett E.,
Yap P., Simmonds P.;
RT "Analysis of a new hepatitis C virus type and its phylogenetic
relationship to existing variants.";
RL J. Gen. Virol. 73:1131-1141(1992).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
virus: identification of a new virus type and restrictions on sequence
diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; D17763; BAA04609.1; -; Genomic_RNA.
DR PIR; PQ0401; PQ0401.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR SMT; Q81258; 1035-1663, 2431-2996.

```

```

DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR002518; Pept_U39 HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002016; Peroxidase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
KW Polypeptide.
FT CHAIN 1 191
FT CHAIN 192 383
FT CHAIN 384 735
FT CHAIN 736 1012
FT CHAIN 1013 1663
FT CHAIN 1664 1717
FT CHAIN 1718 1978
FT CHAIN 1979 2430
FT CHAIN 2431 3021
SQ SEQUENCE 3021 AA; 329578 MW; 38712CCBC0C19562 CRC64;

Query Match 96.9%; Score 597; DB 2; Length 3021;
Best Local Similarity 96.6%; Pred. No. 1.7e-50;
Matches 114; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQQYD 60
Db 1653 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGVVIGHIELGKPAIVDPKVELYQQYD 1712

Qy 61 EMESCSQAAPYIEQAQVIAHQFKGVLLGLLQRAQQQAQVIEPIVTTNNQKLEAFWHKH 118
Db 1713 EMESCSQAAPYIEQAQVIAHQFKGVLLGLLQRAQQQAQVIEPIVTTNNQKLEAFWHKH 1770

RESULT 5
Q81595_9HEPC
ID Q81595_9HEPC PRELIMINARY; PRT; 133 AA.
AC Q81595;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TremBLrel. 24, Last annotation update)
DE Nonstructural protein 4 (Fragment).
GN Name=NS4;

```

```
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93249436; PubMed=7683463;
RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;
RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV
  type 3.";
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).
DR EMBL; D14602; BAA03451.1; -; Genomic_RNA.
DR HSP; P26663; 1CUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 133
FT NON_TER 133
SQ SEQUENCE 133 AA; 14781 MW; 4BFF2128FD301691 CRC64;

Query Match      95.8%; Score 590; DB 2; Length 133;
Best Local Similarity 93.2%; Pred. No. 3.5e-51;
Matches 110; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSGCVVIVGHIELGKPAIVDPKVELVYQQYD 60
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSGCVVIVGHIELGKPAIVDPKVELVYQQYD 75
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLLGLLQRLATQQQAQVIEPIVTTNWQKLEAFWHKH 118
Db 76 EMECSQAAPYIEQAQVIAHQFKGVLLGLLQRLATQQQAQVIEPIVTTNWQKLEAFWHKH 133
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLLGLLQRLATQQQAQVIEPIVTTNWQKLEAFWHKH 118
Db 81 EMECSQAAPYIEQAQVIAHQFKGVLLGLLQRLATQQQAQVIEPIVTTNWQKLEAFWHKH 138

RESULT 6
Q68233_9HEPC
ID Q68233_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68233_9HEPC
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=3a;
RC MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14269; AAC53958.1; -; Genomic_RNA.
DR HSP; P26663; 1CUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
FT NON_TER 138
SQ SEQUENCE 138 AA; 15379 MW; 45236C0E5427B19F CRC64;

Query Match      95.6%; Score 589; DB 2; Length 138;
Best Local Similarity 94.1%; Pred. No. 4.5e-51;
Matches 111; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSGCVVIVGHIELGKPAIVDPKVELVYQQYD 60
Db 21 ACMSADLEVTSTWLLGGVLAALAAAYCLSGCVVIVGHIELGKPAIVDPKVELVYQQYD 80
```

```
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLLGLLQRLATQQQAQVIEPIVTTNWQKLEAFWHKH 118
Db 81 EMECSQAAPYIEQAQVIAHQFKGVLLGLLQRLATQQQAQVIEPIVTTNWQKLEAFWHKH 138

RESULT 7
Q81495_9HEPC
ID Q81495_9HEPC PRELIMINARY; PRT; 3021 AA.
AC Q81495_9HEPC
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=K3a;
RC MEDLINE=95053917; PubMed=7964640;
RA Yamada N., Manihara K., Mizokami M., Ohba K., Takada A., Tetsuami M.,
  Date T.;
RT "Full-length sequence of the genome of hepatitis C virus type 3a:
  comparative study with different genotypes.";
RL J. Gen. Virol. 75:3279-3284(1994).
DR EMBL; D28917; BAA06044.1; -; Genomic_RNA.
DR HSP; P26664; 1HEI.
DR SNR; Q81495; 1035-1663; 2431-2996.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW SEQUENCE 3021 AA; 328387 MW; A97418FF36C062A4 CRC64;

Query Match      94.0%; Score 579; DB 2; Length 3021;
Best Local Similarity 93.2%; Pred. No. 1.1e-48;
Matches 110; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
```

```
Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 1653 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 1712
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
Db 1713 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTVSNQKLEVLWHKH 1770

RESULT 8
Q68239_9HEPC
ID Q68239_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68239;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14275; AAC53964.1; -; Genomic_RNA.
DR HSP; P26663; 1CU1.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15311 MW; B45AC0F8917DBAAC CRC64;

Query Match 93.7%; Score 577; DB 2; Length 138;
Best Local Similarity 93.2%; Pred. No. 7.3e-50;
Matches 110; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 21 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 80
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
Db 81 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVLTWQKLETFWHKH 138

RESULT 9
Q68241_9HEPC
ID Q68241_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68241;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14275; AAC53964.1; -; Genomic_RNA.
DR HSP; P26663; 1CU1.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15311 MW; B45AC0F8917DBAAC CRC64;

Query Match 93.7%; Score 577; DB 2; Length 138;
Best Local Similarity 93.2%; Pred. No. 7.3e-50;
Matches 110; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 21 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 80
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
Db 81 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVLTWQKLETFWHKH 138

RESULT 10
O56637_9HEPC
ID O56637_9HEPC PRELIMINARY; PRT; 193 AA.
AC O56637;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein (Fragment).
GN Name=NS4;
OS Hepatitis C virus type 3g.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=42792;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=type 3g;
RA Panigrahi A.K., Panda S.K.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF042791; AAB97852.1; -; Genomic_RNA.
DR HSP; P26663; 1CU1.
DR MEROPS; S29.001; 1-85.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 21493 MW; AF7B8E3FB3B69505 CRC64;

Query Match 87.5%; Score 539; DB 2; Length 193;
Best Local Similarity 83.9%; Pred. No. 6.8e-46;
Matches 99; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 75 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 134
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
Db 135 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTVSNQKLEAFWHKH 192

RESULT 11
O81487_9HEPC
ID O81487_9HEPC PRELIMINARY; PRT; 3023 AA.
AC O81487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
```

```
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14277; AAC53966.1; -; Genomic_RNA.
DR HSP; P26663; 1CU1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15277 MW; 275F2F928F5A10E4 CRC64;

Query Match 92.9%; Score 572; DB 2; Length 138;
Best Local Similarity 92.4%; Pred. No. 2.3e-49;
Matches 109; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 21 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 80
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
Db 81 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVLTWQKLETFWHKH 138

RESULT 10
O56637_9HEPC
ID O56637_9HEPC PRELIMINARY; PRT; 193 AA.
AC O56637;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein (Fragment).
GN Name=NS4;
OS Hepatitis C virus type 3g.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=42792;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=type 3g;
RA Panigrahi A.K., Panda S.K.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF042791; AAB97852.1; -; Genomic_RNA.
DR HSP; P26663; 1CU1.
DR MEROPS; S29.001; 1-85.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 193
SQ SEQUENCE 193 AA; 21493 MW; AF7B8E3FB3B69505 CRC64;

Query Match 87.5%; Score 539; DB 2; Length 193;
Best Local Similarity 83.9%; Pred. No. 6.8e-46;
Matches 99; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 60
Db 75 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPAIVDPKEVLYQQYD 134
Qy 61 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTTNNQKLEAFWHKH 118
Db 135 EMECSQAAPYIEQAQVIAHQFKGVLGGLLQORATQQQAVIEPIVTVSNQKLEAFWHKH 192

RESULT 11
O81487_9HEPC
ID O81487_9HEPC PRELIMINARY; PRT; 3023 AA.
AC O81487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
```



```
Oy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIYVPDKKVLVYQYDE 61
Db 1653 CMAADLEVTATSAWLLGGVMAALTAAYCLSVGSGWIVGHVLLGGKPAIYVPDKKVLVYQYDE 1712
Oy 62 MEECSQAAPYIEQAQVIAHQFQKGVLLGLLQATQQAAVIEPIVTTNWOKLEAFWFKH 118
Db 1713 MEECSRAAPYIEQAQVIAHQFQKGVLLGLLQADQAAIKPIATPYWKLTFNSKH 1769

RESULT 13
Q68223_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68223_9HEPC PRELIMINARY; PRT; 138 AA.
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14259; AAC53948.1; -; Genomic_RNA.
DR HSPF; P27958; IHEI.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15392 MW; F8CB866A53AA907B CRC64;

Query Match 76.8%; Score 473; DB 2; Length 138;
Best Local Similarity 75.2%; Pred. No. 2.1e-39;
Matches 88; Conservative 12; Mismatches 17; Indels 0; Gaps 0;

Oy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIYVPDKKVLVYQYDE 61
Db 22 CMSADLEVTSTWLLGGVLAALAAAYCLSTGCVVIVGRVLSGKPAIIPDREVLVYREFDE 81
Oy 62 MEECSQAAPYIEQAQVIAHQFQKGVLLGLLQATQQAAVIEPIVTTNWOKLEAFWFKH 118
Db 82 MEECSQHLPHYIEQGMMLAEQFKQKALGLLQATASQAEVITPVVQTNWOKLEAFWAKH 138

RESULT 14
Q68K66_9HEPC PRELIMINARY; PRT; 658 AA.
AC Q68K66_9HEPC PRELIMINARY; PRT; 658 AA.
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in HCV/HIV Coinfected Subjects.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685564; AAT94247.1; -; Genomic_RNA.
```

```
DR SMR; Q68K66; 3-551.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 658
SQ SEQUENCE 658 AA; 70676 MW; B1C0F9F7BFBAE4E3 CRC64;

Query Match 76.3%; Score 470; DB 2; Length 658;
Best Local Similarity 74.4%; Pred. No. 2.1e-38;
Matches 87; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

Oy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIYVPDKKVLVYQYDE 61
Db 542 CMSADLEVTSTWLLGGVLAALAAAYCLSTGCVVIVGRVLSGKPAIIPDREVLVYREFDE 601
Oy 62 MEECSQAAPYIEQAQVIAHQFQKGVLLGLLQATQQAAVIEPIVTTNWOKLEAFWFKH 118
Db 602 MEECSQHLPHYIEQGMMLAEQFKQKALGLLQATASQAEVITPVVQTNWOKLEAFWAKH 658

RESULT 15
Q68224_9HEPC PRELIMINARY; PRT; 138 AA.
ID Q68224_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68224_9HEPC PRELIMINARY; PRT; 138 AA.
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14260; AAC53949.1; -; Genomic_RNA.
DR HSPF; P27958; IHEI.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15295 MW; 01335FB49A841A53 CRC64;

Query Match 75.8%; Score 467; DB 2; Length 138;
Best Local Similarity 75.2%; Pred. No. 8.4e-39;
Matches 88; Conservative 10; Mismatches 19; Indels 0; Gaps 0;

Oy 2 CMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIYVPDKKVLVYQYDE 61
```

Db	22	CMSADLEVTSTWLVGGVLAALAAAYCLSTGCVIVGRIVLSGKPAIIPDREVLREFDE	81
Qy	62	MECSCQAPYIEQAOVIAHQFKGKVLGLQRATQQQAVIEBPVITTNWQKLEAFWKKH	118
Db	82	MECSCQHLPYIEQGMILAEOFKQKALGLLOTASRQAEVITPAVQTNWQKLEAFWAKH	138

Search completed: January 27, 2006, 19:13:26
Job time : 163 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:15:20 ; Search time 135 Seconds
(without alignments)
384.050 Million cell updates/sec

Title: US-09-638-693a-36_COPY_16_133

Perfect score: 118

Sequence: 1 ACMSADLEVTSTWLLGGV.....VIEPIVTNNQKLEAFWHKH 118

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2443163 seqs, 439378781 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 2000 summaries

Database : A_Geneseq_21:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003s:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*
- 9: Geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	118	100.0	133	2	Aar63288 Polypepti
2	83	70.3	128	2	Aar37937 HCV NS4 r
3	83	70.3	128	2	Aar37932 HCV NS4 r
4	83	70.3	133	2	Aar63289 Polypepti
5	83	70.3	133	2	Aar63290 Polypepti
6	83	70.3	209	2	Aar63390 HCV polyp
7	47	39.8	117	2	Aar37934 HCV NS4 r
8	47	39.8	117	2	Aar37935 HCV NS4 r
9	47	39.8	127	2	Aar37936 HCV NS4 r
10	47	39.8	127	2	Aar37933 HCV NS4 r
11	47	39.8	133	2	Aar63287 Polypepti
12	47	39.8	133	2	Aar63286 Polypepti
13	47	39.8	3021	9	Adx40823 HCV polyp
14	47	39.8	3021	9	Adx40824 HCV polyp
15	42	35.6	3021	9	Adx40820 HCV polyp
16	41	34.7	829	5	AAE18690 Multiple
17	41	34.7	829	5	ADC06769 Chimeric
18	41	34.7	829	8	ADL66807 HCV multi
19	41	34.7	1099	5	AAU76378 HCV multi
20	41	34.7	1099	6	ABG72262 HCV multi
21	41	34.7	1099	8	ADL66809 HCV multi
22	30	25.4	3023	2	AAR94462 Hepatitis
23	30	25.4	3023	9	Adx40803 HCV polyp
24	23	19.5	3021	9	Adx40813 HCV polyp

Aar63316	Peptide f	20	16.9	25
Aar41164	HCV (type	22	16.9	26
Aar37941	HCV NS-4	19	16.1	27
Adx40798	HCV polym	3018	16.1	28
Aar63318	Peptide f	20	15.3	29
Aar41165	HCV (type	22	15.3	30
Aay06672	Amino aci	352	15.3	31
Adx40827	HCV polym	3016	15.3	32
Adx40810	HCV polym	3019	15.3	33
Adx40780	HCV polym	3008	15.3	34
Adx40825	HCV polym	3015	12.7	35
Aar63316	Peptide f	20	12.7	36
Aar37938	HCV NS-4	18	11.9	37
Aay67804	Peptide #	35	11.9	38
Adv23000	HCV H77 i	18	11.0	39
Aar63317	Peptide f	20	11.0	40
Aar63319	Peptide f	20	11.0	41
Aar41166	HCV (type	24	11.0	42
Aau84709	HCV HepC1	30	11.0	43
Aau84708	HCV HepC1	30	11.0	44
Aau09248	HCV trunc	51	11.0	45
Abb77267	HCV bait	51	11.0	46
Aaw09242	HCV NS4A	54	11.0	47
Aaw04579	Hepatitis	54	11.0	48
Aaw01651	HCV NS4A	54	11.0	49
Aaw47144	HCV NS4A	54	11.0	50
Aay57203	HCV NS4A	54	11.0	51
Aay57207	HCV NS4A	54	11.0	52
AAE13186	Hepatitis	54	11.0	53
AAE10069	Hepatitis	54	11.0	54
AAE21848	Hepatitis	54	11.0	55
AAE19894	Hepatitis	54	11.0	56
AAE19906	Hepatitis	54	11.0	57
ABW00345	Hepatitis	54	11.0	58
ABW00357	Hepatitis	54	11.0	59
ADG47670	HCV NS3/4	54	11.0	60
AAp92020	Sequence	135	11.0	61
AAp90137	Sequence	135	11.0	62
Aaw09240	HCV insol	234	11.0	63
AAW01649	HCV NS3 p	234	11.0	64
AAp92021	Polyepti	237	11.0	65
AAp90138	Peptide e	237	11.0	66
AAW09238	HCV solub	237	11.0	67
AAW01647	HCV NS3 s	237	11.0	68
AAW09239	HCV solub	250	11.0	69
AAW01648	HCV NS3 s	250	11.0	70
ADT77851	Hepatitis	252	11.0	71
AAW09237	HCV solub	255	11.0	72
AAW04572	HCV NS3-N	255	11.0	73
AAW01646	HCV NS3 s	255	11.0	74
AAW47147	Chloramph	255	11.0	75
AAy57206	Amino aci	255	11.0	76
AAW71272	Amino aci	258	11.0	77
AAW09236	HCV solub	270	11.0	78
AAW04571	HCV NS3-N	270	11.0	79
AAW01645	HCV NS3 s	270	11.0	80
AAAR30064	HCV NS3/N	313	11.0	81
AAy22022	HCV E1 pe	313	11.0	82
AAW75483	Hepatitis	313	11.0	83
AAAR33557	Antigen p	342	11.0	84
ABW69013	HCV recom	342	11.0	85
ABW01878	HCV CKS-C	342	11.0	86
AAAB69014	HCV recom	344	11.0	87
ABW01879	HCV CKS-C	344	11.0	88
AAAR33559	Antigen p	352	11.0	89
AAAB69015	HCV recom	352	11.0	90
ABW01880	HCV CKS-C	352	11.0	91
AAAR33560	Antigen p	357	11.0	92
ABW69016	HCV recom	357	11.0	93
ABW01881	HCV CKS-C	357	11.0	94
AAAR33561	Antigen p	362	11.0	95
ABW69017	HCV recom	362	11.0	96
ABW01882	HCV CKS-C	362	11.0	97

98	13	11.0	363	2	AAR23999	Aar23999	Open read	171	13	11.0	686	7	ABW00376	Abw00376	Hepatitis
99	13	11.0	363	2	AAR33563	Aar33563	Antigen p	172	13	11.0	686	8	ADG47659	Adg47659	HCV NS3/4
100	13	11.0	363	2	AAR90933	Aar90933	HCV antigen	173	13	11.0	686	8	ADG47665	Adg47665	HCV NS3/4
101	13	11.0	363	2	AAB69019	Aab69019	HCV recom	174	13	11.0	686	8	ADG47666	Adg47666	HCV NS3/4
102	13	11.0	363	7	ABW01884	Abw01884	HCV CKS-C	175	13	11.0	686	8	ADG47662	Adg47662	HCV NS3/4
103	13	11.0	364	2	AAR33564	Aar33564	Antigen p	176	13	11.0	686	8	ADG47693	Adg47693	HCV NS3/4
104	13	11.0	364	4	AAB69020	Abb69020	HCV recom	177	13	11.0	686	8	ADG47660	Adg47660	HCV NS3/4
105	13	11.0	364	7	ABW01885	Abw01885	HCV CKS-C	178	13	11.0	686	8	ADG47661	Adg47661	HCV NS3/4
106	13	11.0	365	2	AAR33562	Aar33562	Antigen p	179	13	11.0	686	8	ADG47667	Adg47667	HCV NS3/4
107	13	11.0	365	4	AAB69018	Aab69018	HCV recom	180	13	11.0	686	8	ADG47664	Adg47664	HCV NS3/4
108	13	11.0	365	7	ABW01883	Abw01883	HCV CKS-C	181	13	11.0	686	8	ADG47663	Adg47663	HCV NS3/4
109	13	11.0	382	1	AAP92048	Aap92048	Carboxy-t	182	13	11.0	686	8	ADG47668	Adg47668	HCV NS3/4
110	13	11.0	382	1	AAP90182	Aap90182	C terminu	183	13	11.0	686	8	ADL66805	Adl66805	HCV NS3/4
111	13	11.0	460	1	AAP92024	Aap92024	Polypepti	184	13	11.0	728	5	AEL18688	Ael18688	NS3/4a mu
112	13	11.0	460	1	AAP90141	Aap90141	Protein s	185	13	11.0	728	7	ADC06766	Adc06766	HCV mutan
113	13	11.0	504	2	AAY01622	Aay01622	Protein s	186	13	11.0	858	1	AAP90146	Aap90146	ORF exten
114	13	11.0	507	2	AAR66631	Aar66631	HCV J1 NS	187	13	11.0	859	1	AAP92029	Aap92029	HCV prote
115	13	11.0	592	2	AAR33565	Aar33565	CKS-HCV a	188	13	11.0	971	2	AAR33569	Aar33569	CKS-HCV a
116	13	11.0	592	4	AAB69023	Abb69023	HCV recom	189	13	11.0	971	4	AAB69027	Abb69027	HCV recom
117	13	11.0	592	7	ABW01888	Abw01888	HCV CKS-C	190	13	11.0	971	7	ABW01892	Abw01892	HCV CKS-C
118	13	11.0	594	2	AAR33566	Aar33566	CKS-HCV a	191	13	11.0	973	2	AAR33570	Aar33570	CKS-HCV a
119	13	11.0	594	4	AAB69024	Abb69024	HCV recom	192	13	11.0	973	4	AAB69028	Abb69028	HCV recom
120	13	11.0	594	7	ABW01889	Abw01889	HCV CKS-C	193	13	11.0	973	7	ABW01893	Abw01893	CKS-C
121	13	11.0	597	2	AAR21571	Aar21571	HCV CKS-C	194	13	11.0	992	2	AAR33571	Aar33571	CKS-HCV a
122	13	11.0	597	2	AAR33638	Aar33638	HCV C100D	195	13	11.0	992	2	AAB69029	Abb69029	HCV recom
123	13	11.0	597	2	AAR33630	Aar33630	HCV C100D	196	13	11.0	992	7	ABW01894	Abw01894	HCV CKS-C
124	13	11.0	597	2	AAR33580	Aar33580	HCV-C100D	197	13	11.0	1021	2	AAM40481	Aaw34481	HCV antigen
125	13	11.0	597	4	AAB51378	Aab51378	HCV recom	198	13	11.0	1021	2	AAM40039	Aaw40039	Fusion pr
126	13	11.0	597	7	ABW01864	Abw01864	HCV CKS-C	199	13	11.0	1021	5	AEE22050	Aae22050	psOD/c200
127	13	11.0	599	2	AAR21572	Aar21572	HCV CKS-C	200	13	11.0	1210	9	ADY84770	Ady84770	HCV prote
128	13	11.0	599	2	AAR33639	Aar33639	HCV C100D	201	13	11.0	1210	9	ADY84768	Ady84768	HCV prote
129	13	11.0	599	2	AAR33601	Aar33601	HCV C100D	202	13	11.0	1766	1	AAP92041	Aap92041	Hepatitis
130	13	11.0	599	2	AAR33581	Aar33581	HCV-C100D	203	13	11.0	1771	4	AAB62631	Aab62631	HCV NS35
131	13	11.0	599	4	AAB51379	Aab51379	HCV recom	204	13	11.0	1771	4	AAB62634	Aab62634	Amino aci
132	13	11.0	599	7	ABW01865	Abw01865	HCV-CKS-C	205	13	11.0	1771	4	AAB62635	Aab62635	Amino aci
133	13	11.0	613	2	AAR33567	Aar33567	CKS-HCV a	206	13	11.0	1771	4	AAB62632	Aab62632	HCV delNS
134	13	11.0	613	4	AAB69025	Abb69025	HCV recom	207	13	11.0	1892	4	AAB62636	Aab62636	Amino aci
135	13	11.0	613	7	ABW01890	Abw01890	HCV CKS-C	208	13	11.0	1892	8	AD134636	Adi34636	HCVmodifi
136	13	11.0	685	2	AAR71271	Aar71271	Protein e	209	13	11.0	1892	8	ADO00774	Ado00774	HCV NS345
137	13	11.0	686	4	AAB62633	Aab62633	HCV NG34A	210	13	11.0	1911	4	AAB62638	Aab62638	Amino aci
138	13	11.0	686	5	AAU76377	Aau76377	Hepatitis	211	13	11.0	1921	4	AAB62639	Aab62639	Amino aci
139	13	11.0	686	5	AAU76377	Aau76377	Hepatitis	212	13	11.0	1944	4	AAB62637	Aab62637	Amino aci
140	13	11.0	686	5	AAB21839	Aab21839	Hepatitis	213	13	11.0	1986	8	ADR38451	Adr38451	Hepatitis
141	13	11.0	686	5	AAB21845	Aab21845	Hepatitis	214	13	11.0	1986	9	AEA62092	Aea62092	Polyprote
142	13	11.0	686	5	AAB21842	Aab21842	Hepatitis	215	13	11.0	1987	5	AAU84802	Aau84802	HCV HepC
143	13	11.0	686	5	AAB21837	Aab21837	Hepatitis	216	13	11.0	2011	5	AAU84800	Aau84800	HCV HepC
144	13	11.0	686	5	AAB21838	Aab21838	Hepatitis	217	13	11.0	2202	6	AAO26783	Aao26783	Protein d
145	13	11.0	686	5	AAB21840	Aab21840	Hepatitis	218	13	11.0	2261	1	AAP90164	Aap90164	Peptide e
146	13	11.0	686	5	AAB21846	Aab21846	Hepatitis	219	13	11.0	2301	1	AAP92047	Aap92047	HCV prote
147	13	11.0	686	5	AAB21841	Aab21841	Hepatitis	220	13	11.0	2435	2	AAR25135	Aar25135	HCV polyp
148	13	11.0	686	5	AAB21843	Aab21843	Hepatitis	221	13	11.0	2436	1	AAP92050	Aap92050	HCV prote
149	13	11.0	686	5	AAB21844	Aab21844	Hepatitis	222	13	11.0	2436	1	AAP90288	Aap90288	Peptide e
150	13	11.0	686	5	AAE19908	Aae19908	Hepatitis	223	13	11.0	2436	2	AAR28582	Aar28582	HCV amino
151	13	11.0	686	5	AAE19919	Aae19919	Hepatitis	224	13	11.0	2631	6	AAO26785	Aao26785	Protein d
152	13	11.0	686	5	AAE19922	Aae19922	Hepatitis	225	13	11.0	2772	3	AAR08123	Aar08123	Hepatitis
153	13	11.0	686	5	AAE19907	Aae19907	Hepatitis	226	13	11.0	2772	3	AAB18540	Aab18540	Protein e
154	13	11.0	686	5	AAE19924	Aae19924	Hepatitis	227	13	11.0	2772	8	ADN35976	Adn35976	HCV cdNA
155	13	11.0	686	5	AAE19920	Aae19920	Hepatitis	228	13	11.0	2816	2	AAR34009	Aar34009	HCV-1 pol
156	13	11.0	686	5	AAE19921	Aae19921	Hepatitis	229	13	11.0	2894	2	AAR24440	Aar24440	Composite
157	13	11.0	686	5	AAE19900	Aae19900	Hepatitis	230	13	11.0	2894	2	AAR70230	Aar70230	Composite
158	13	11.0	686	5	AAE19923	Aae19923	Hepatitis	231	13	11.0	2955	2	AAR08124	Aar08124	Hepatitis
159	13	11.0	686	5	AAE19925	Aae19925	Hepatitis	232	13	11.0	2955	2	AAI14975	Aay14975	Amino aci
160	13	11.0	686	6	ABG72261	Abg72261	HCV-1 NS3	233	13	11.0	2955	3	AAB18541	Aab18541	Polyprote
161	13	11.0	686	7	ADC06767	Adc06767	HCV mutan	234	13	11.0	2955	8	ADN35978	Adn35978	HCV cdNA
162	13	11.0	686	7	ABW00374	Abw00374	Hepatitis	235	13	11.0	2984	4	AAE00449	Aae00449	Hepatitis
163	13	11.0	686	7	ABW00375	Abw00375	Hepatitis	236	13	11.0	2984	4	AAE00447	Aae00447	Hepatitis
164	13	11.0	686	7	ABW00370	Abw00370	Hepatitis	237	13	11.0	2984	4	AAE00442	Aae00442	Hepatitis
165	13	11.0	686	7	ABW00371	Abw00371	Hepatitis	238	13	11.0	3002	7	ADM24822	Adm24822	Hepatitis
166	13	11.0	686	7	ABW00359	Abw00359	Hepatitis	239	13	11.0	3011	2	AAR21519	Aar21519	Compiled
167	13	11.0	686	7	ABW00351	Abw00351	Hepatitis	240	13	11.0	3011	2	AAR22154	Aar22154	NANBV Hut
168	13	11.0	686	7	ABW00373	Abw00373	Hepatitis	241	13	11.0	3011	2	AAR31621	Aar31621	Hepatitis
169	13	11.0	686	7	ABW00358	Abw00358	Hepatitis	242	13	11.0	3011	2	AAR40119	Aar40119	HCV genom
170	13	11.0	686	7	ABW00372	Abw00372	Hepatitis	243	13	11.0	3011	2	AAR40120	Aar40120	HCV genom

244	13	11.0	3011	2	AAR66995	Hepatitis	317	12	10.2	477	2	AAR29866	Aar29866 HCV NS2-N
245	13	11.0	3011	2	AAR79232	HCV seque	318	12	10.2	477	2	AAR29867	Aar29867 HCV NS2-N
246	13	11.0	3011	2	AAR67588	Hepatitis	319	12	10.2	477	2	AAR29865	Aar29865 HCV NS2-N
247	13	11.0	3011	2	AAR90931	Hepatitis	320	12	10.2	685	9	ADZ99621	Adz99621 Hepatitis
248	13	11.0	3011	2	AAW34480	HCV poly	321	12	10.2	685	9	ADZ99616	Adz99616 Hepatitis
249	13	11.0	3011	2	AAW77397	Hepatitis	322	12	10.2	685	9	ADZ99623	Adz99623 Hepatitis
250	13	11.0	3011	2	AAW77398	Hepatitis	323	12	10.2	685	9	ADZ99622	Adz99622 Hepatitis
251	13	11.0	3011	2	AAW40038	HCV poly	324	12	10.2	685	9	ADZ99620	Adz99620 Hepatitis
252	13	11.0	3011	2	AAW98021	Infectiou	325	12	10.2	697	8	ADL17782	Adl17782 Hepatitis
253	13	11.0	3011	2	AAW98020	Infectiou	326	12	10.2	697	8	ADL17782	Adl17782 Hepatitis
254	13	11.0	3011	4	AAW59173	Protein e	327	12	10.2	767	2	AAR80044	Aar80044 Hepatitis
255	13	11.0	3011	4	AAW59173	Protein e	328	12	10.2	767	2	AAR80044	Aar80044 Hepatitis
256	13	11.0	3011	5	AAU99290	Hepatitis	329	12	10.2	768	2	AAR29868	Aar29868 HCV NS2-N
257	13	11.0	3011	5	AAU99290	Hepatitis	330	12	10.2	768	2	AAR29868	Aar29868 HCV NS2-N
258	13	11.0	3011	5	AAU84597	Hepatitis	331	12	10.2	768	2	AAR29868	Aar29868 HCV NS2-N
259	13	11.0	3011	5	AAE22049	Hepatitis	332	12	10.2	768	2	AAR29868	Aar29868 HCV NS2-N
260	13	11.0	3011	5	AAE22052	Hepatitis	333	12	10.2	916	2	AAR82693	Aar82693 HCV parti
261	13	11.0	3011	5	AAU79221	Hepatitis	334	12	10.2	923	2	AAR82696	Aar82696 HCV parti
262	13	11.0	3011	5	AAE19888	Hepatitis	335	12	10.2	947	9	ADV66342	Adv66342 Hepatitis
263	13	11.0	3011	6	ABP71460	Protein d	336	12	10.2	1188	2	AAR29660	Aar29660 HCV NS2-N
264	13	11.0	3011	6	ABU61849	HCV-H. 8/	337	12	10.2	1188	2	AAR29660	Aar29660 HCV NS2-N
265	13	11.0	3011	7	ABW00339	Hepatitis	338	12	10.2	1250	2	AAR12599	Aar12599 Portion O
266	13	11.0	3011	8	ADH79949	E2 HCV en	339	12	10.2	1394	9	AEA32849	Aea32849 Modified
267	13	11.0	3011	8	ADH79949	E2 HCV en	340	12	10.2	1411	2	AAR29533	Aar29533 HCV NS4-N
268	13	11.0	3011	8	ADL23107	Hepatitis	341	12	10.2	1736	4	AAAB36932	Aab36932 Hepatitis
269	13	11.0	3011	8	ADJ64256	Hepatitis	342	12	10.2	1985	5	AAO18001	Aao18001 Hepatitis
270	13	11.0	3011	8	ADL72983	Hepatitis	343	12	10.2	1985	5	AAE15729	Aae15729 Hepatitis
271	13	11.0	3011	8	ADR29357	Hepatitis	344	12	10.2	1985	5	AAE15731	Aae15731 Hepatitis
272	13	11.0	3011	8	ADH79949	E2 HCV en	345	12	10.2	1985	5	AAE15720	Aae15720 Hepatitis
273	13	11.0	3011	9	ADH79949	E2 HCV en	346	12	10.2	1985	5	AAE15727	Aae15727 Hepatitis
274	13	11.0	3011	9	ADH79949	E2 HCV en	347	12	10.2	1985	5	AAE15728	Aae15728 Hepatitis
275	13	11.0	3011	9	ADH79949	E2 HCV en	348	12	10.2	1985	5	AAE15722	Aae15722 Hepatitis
276	13	11.0	3011	9	ADH79949	E2 HCV en	349	12	10.2	1985	5	AAE15730	Aae15730 Hepatitis
277	13	11.0	3011	9	ADH79949	E2 HCV en	350	12	10.2	1985	6	ABU09574	Abu09574 HCV Met-N
278	13	11.0	3011	9	ADH79949	E2 HCV en	351	12	10.2	1985	6	ABU09575	Abu09575 HCV Met-N
279	13	11.0	3011	9	ADH79949	E2 HCV en	352	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
280	13	11.0	3011	9	ADH79949	E2 HCV en	353	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
281	13	11.0	3011	9	ADH79949	E2 HCV en	354	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
282	13	11.0	3011	9	ADH79949	E2 HCV en	355	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
283	13	11.0	3011	9	ADH79949	E2 HCV en	356	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
284	13	11.0	3011	9	ADH79949	E2 HCV en	357	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
285	13	11.0	3011	9	ADH79949	E2 HCV en	358	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
286	13	11.0	3011	9	ADH79949	E2 HCV en	359	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
287	13	11.0	3011	9	ADH79949	E2 HCV en	360	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
288	13	11.0	3011	9	ADH79949	E2 HCV en	361	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
289	13	11.0	3011	9	ADH79949	E2 HCV en	362	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
290	13	11.0	3011	9	ADH79949	E2 HCV en	363	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
291	13	11.0	3011	9	ADH79949	E2 HCV en	364	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
292	13	11.0	3011	9	ADH79949	E2 HCV en	365	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
293	13	11.0	3011	9	ADH79949	E2 HCV en	366	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
294	13	11.0	3011	9	ADH79949	E2 HCV en	367	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
295	13	11.0	3011	9	ADH79949	E2 HCV en	368	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
296	13	11.0	3011	9	ADH79949	E2 HCV en	369	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
297	13	11.0	3011	9	ADH79949	E2 HCV en	370	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
298	13	11.0	3011	9	ADH79949	E2 HCV en	371	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
299	13	11.0	3011	9	ADH79949	E2 HCV en	372	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
300	13	11.0	3011	9	ADH79949	E2 HCV en	373	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
301	13	11.0	3011	9	ADH79949	E2 HCV en	374	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
302	13	11.0	3011	9	ADH79949	E2 HCV en	375	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
303	13	11.0	3011	9	ADH79949	E2 HCV en	376	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
304	13	11.0	3011	9	ADH79949	E2 HCV en	377	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
305	13	11.0	3011	9	ADH79949	E2 HCV en	378	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
306	13	11.0	3011	9	ADH79949	E2 HCV en	379	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
307	13	11.0	3011	9	ADH79949	E2 HCV en	380	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
308	13	11.0	3011	9	ADH79949	E2 HCV en	381	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
309	13	11.0	3011	9	ADH79949	E2 HCV en	382	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
310	13	11.0	3011	9	ADH79949	E2 HCV en	383	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
311	13	11.0	3011	9	ADH79949	E2 HCV en	384	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
312	13	11.0	3011	9	ADH79949	E2 HCV en	385	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
313	13	11.0	3011	9	ADH79949	E2 HCV en	386	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
314	13	11.0	3011	9	ADH79949	E2 HCV en	387	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
315	13	11.0	3011	9	ADH79949	E2 HCV en	388	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli
316	13	11.0	3011	9	ADH79949	E2 HCV en	389	12	10.2	1985	8	ADJ57846	Adj57846 HCV repli

390	12	10.2	3010	2	AAR68864	Aar68864	Hepatitis	463	10	8.5	15	4	AAJ04008	Hepatitis
391	12	10.2	3010	2	AAR82694	Aar82694	Partial H	464	10	8.5	15	4	AAJ03473	Hepatitis
392	12	10.2	3010	2	AAJ06423	Aay06423	Non-A, no	465	10	8.5	15	7	ADW36129	HLA bindi
393	12	10.2	3010	2	AAW98022	Aaw98022	Infectiou	466	10	8.5	15	7	ADW35960	ADW35960
394	12	10.2	3010	4	AAB59174	Aab59174	Protein e	467	10	8.5	15	9	ADV23001	HCV H77 1
395	12	10.2	3010	4	AAB31170	Aab31170	Amino aci	468	10	8.5	303	2	AAR20713	C10-18 NA
396	12	10.2	3010	5	ABG32458	Abg32458	Hepatitis	469	9	7.6	9	2	AAV46886	Immunogen
397	12	10.2	3010	5	ABG32459	Abg32459	Hepatitis	470	9	7.6	9	4	AAJ00391	Hepatitis
398	12	10.2	3010	5	ABG32451	Abg32451	Hepatitis	471	9	7.6	9	4	AAJ00105	Hepatitis
399	12	10.2	3010	5	ABG32455	Abg32455	Hepatitis	472	9	7.6	9	4	AAJ00900	Hepatitis
400	12	10.2	3010	5	ABG32457	Abg32457	Hepatitis	473	9	7.6	9	4	AAJ02704	Hepatitis
401	12	10.2	3010	5	ABG32460	Abg32460	Hepatitis	474	9	7.6	9	4	AAJ01418	Hepatitis
402	12	10.2	3010	5	ABG32453	Abg32453	Hepatitis	475	9	7.6	9	4	AAJ03201	Hepatitis
403	12	10.2	3010	5	ABG32461	Abg32461	Hepatitis	476	9	7.6	9	4	AAJ02222	Hepatitis
404	12	10.2	3010	5	ABG32454	Abg32454	Hepatitis	477	9	7.6	9	4	AAJ03317	Hepatitis
405	12	10.2	3010	5	ABG32452	Abg32452	Hepatitis	478	9	7.6	9	4	AAJ03453	Hepatitis
406	12	10.2	3010	7	ADF88597	Adf88597	Hepatitis	479	9	7.6	9	4	AAJ02191	Hepatitis
407	12	10.2	3010	8	ADO36227	Ado36227	Hepatitis	480	9	7.6	9	4	AAJ03012	Hepatitis
408	12	10.2	3010	8	ADO79401	Ado79401	Hepatitis	481	9	7.6	9	4	AAJ01767	Hepatitis
409	12	10.2	3010	9	ADX40800	Adx40800	HCV polym	482	9	7.6	10	4	AAJ02553	Hepatitis
410	12	10.2	3010	9	ADX40801	Adx40801	HCV polym	483	9	7.6	11	4	AAJ02352	Hepatitis
411	12	10.2	3010	9	ADX40811	Adx40811	HCV polym	484	9	7.6	11	4	AAJ00438	Hepatitis
412	12	10.2	3010	9	ADX40795	Adx40795	HCV polym	485	9	7.6	11	4	AAJ01989	Hepatitis
413	12	10.2	3010	9	ADX40796	Adx40796	HCV polym	486	9	7.6	13	4	AAB96868	Hepatitis
414	12	10.2	3010	9	ADX40808	Adx40808	HCV polym	487	9	7.6	15	4	AAJ03558	Hepatitis
415	12	10.2	3010	9	ADX40818	Adx40818	HCV polym	488	9	7.6	15	4	AAJ03705	Hepatitis
416	12	10.2	3010	9	ADX40779	Adx40779	HCV polym	489	9	7.6	15	4	AAJ03253	Hepatitis
417	12	10.2	3010	9	ADX40786	Adx40786	HCV polym	490	9	7.6	15	4	AAJ03643	Hepatitis
418	12	10.2	3010	9	ADX40817	Adx40817	HCV polym	491	9	7.6	18	5	ABB77268	HCV bait
419	12	10.2	3010	9	ADX40794	Adx40794	HCV polym	492	9	7.6	19	5	AAE21994	Hepatitis
420	12	10.2	3010	9	ADX40807	Adx40807	HCV polym	493	9	7.6	31	2	AAW13248	Hepatitis
421	12	10.2	3010	9	ADX40791	Adx40791	HCV polym	494	9	7.6	35	3	AAV67802	Peptide #
422	12	10.2	3010	9	ADX40792	Adx40792	HCV polym	495	9	7.6	48	2	AAW13247	Hepatitis
423	12	10.2	3010	9	ADX40781	Adx40781	HCV polym	496	9	7.6	119	2	AAR63301	Polypepti
424	12	10.2	3010	9	ADX40783	Adx40783	HCV polym	497	9	7.6	481	2	AAR63436	HCV polyp
425	12	10.2	3010	9	ADX40805	Adx40805	HCV polym	498	9	7.6	489	2	AAR63377	Hepatitis
426	12	10.2	3010	9	ADX40816	Adx40816	HCV polym	499	9	7.6	615	2	AAW37807	Nonstruct
427	12	10.2	3010	9	ADX40788	Adx40788	HCV polym	500	9	7.6	631	2	AAR82854	NS3 serin
428	12	10.2	3010	9	ADX40789	Adx40789	HCV polym	501	9	7.6	631	2	AAW26160	Serine pr
429	12	10.2	3010	9	ADX40804	Adx40804	HCV polym	502	9	7.6	631	2	AAW14354	Hepatitis
430	12	10.2	3010	9	ADX40806	Adx40806	HCV polym	503	9	7.6	631	2	AAV15806	HCV strai
431	12	10.2	3010	9	ADX40812	Adx40812	HCV polym	504	9	7.6	631	7	ADL18158	Hepatitis
432	12	10.2	3011	2	AAR34468	Aar34468	Encoded b	505	9	7.6	632	8	ADO36213	Native HC
433	12	10.2	3011	5	ABG32456	Abg32456	Hepatitis	506	9	7.6	632	8	ADO36213	Hepatitis
434	12	10.2	3013	3	ADG40797	Adg40797	HCV polym	507	9	7.6	632	8	ADO79387	Hepatitis
435	12	10.2	3014	2	AAR35207	Aar35207	Hepatitis	508	9	7.6	646	2	AAV17894	HCV NS4A-
436	12	10.2	3014	2	AAR54099	Aar54099	NANBHV E1	509	9	7.6	646	2	AAV17892	HCV NS4A-
437	12	10.2	3019	9	ADX40822	Adx40822	HCV polym	510	9	7.6	646	2	AAV24950	HCV NS4A-
438	12	10.2	3090	7	ADD67962	Add67962	EMCV inte	511	9	7.6	665	2	AAV24942	HCV NS4A-
439	12	10.2	3091	9	ABE17115	Aeb17115	Hepatitis	512	9	7.6	665	2	AAV24941	HCV NS4A-
440	11	9.3	11	4	AAJ00901	Aaj00901	Hepatitis	513	9	7.6	665	2	AAV24945	HCV NS4A-
441	11	9.3	11	4	AAJ00334	Aaj00334	Hepatitis	514	9	7.6	665	2	AAV24944	HCV NS4A-
442	11	9.3	11	4	AAJ01961	Aaj01961	Hepatitis	515	9	7.6	665	2	AAV24940	HCV NS4A-
443	11	9.3	15	4	AAJ03502	Aaj03502	Hepatitis	516	9	7.6	665	2	AAV24943	HCV NS4A-
444	11	9.3	15	4	AAJ03195	Aaj03195	Hepatitis	517	9	7.6	665	2	AAV24947	HCV NS4A-
445	11	9.3	20	2	AAR63320	Aar63320	Peptide f	518	9	7.6	665	2	AAV24946	HCV NS4A-
446	11	9.3	35	3	AAV67803	Aay67803	Peptide #	519	9	7.6	665	6	ADAL12164	Hepatitis
447	11	9.3	3010	5	AAE20477	Aae20477	HCV-S1 fu	520	9	7.6	665	6	ADAL12166	Hepatitis
448	11	9.3	3022	9	ADX40809	Adx40809	HCV polym	521	9	7.6	665	6	ADAL12166	Hepatitis
449	10	8.5	10	2	AAR73113	Aar73113	Antigen f	522	9	7.6	667	2	AAV17893	HCV NS4A-
450	10	8.5	10	2	AAR61533	Aar61533	Peptide f	523	9	7.6	667	2	AAV17891	HCV NS4A-
451	10	8.5	10	2	AAR78955	Aar78955	HCV NS4 1	524	9	7.6	671	2	AAV24949	HCV NS4A-
452	10	8.5	10	4	AAJ02685	Aaj02685	Hepatitis	525	9	7.6	671	2	AAV24948	HCV NS4A-
453	10	8.5	10	4	AAJ01062	Aaj01062	Hepatitis	526	9	7.6	798	8	ADT88222	Hepatitis
454	10	8.5	10	4	AAJ00646	Aaj00646	Hepatitis	527	9	7.6	3014	9	ADX40799	HCV polym
455	10	8.5	10	4	AAJ02192	Aaj02192	Hepatitis	528	9	7.6	3014	9	ADX40821	HCV polym
456	10	8.5	10	4	AAJ01923	Aaj01923	Hepatitis	529	8	6.8	8	4	AAJ02221	Hepatitis
457	10	8.5	10	8	ADE97764	Ade97764	Immunogen	530	8	6.8	8	4	AAJ02562	Hepatitis
458	10	8.5	10	8	ADR11678	Adr11678	HLA-A2.1	531	8	6.8	8	4	AAJ02903	Hepatitis
459	10	8.5	11	4	AAJ02897	Aaj02897	Hepatitis	532	8	6.8	8	4	AAJ00333	Hepatitis
460	10	8.5	11	4	AAJ02554	Aaj02554	Hepatitis	533	8	6.8	8	4	AAJ02190	Hepatitis
461	10	8.5	15	2	AAW85432	Aaw85432	Helper T-	534	8	6.8	8	4	AAJ01634	Hepatitis
462	10	8.5	15	4	AAJ03166	Aaj03166	Hepatitis	535	8	6.8	8	4	AAJ00166	Hepatitis

682	8	6.8	781	2	AAR33574	Aar33574	HCV CKS-3	755	7	5.9	35	5	ABB04543	Abb04543	Hepatitis
683	8	6.8	781	2	AAR52690	Aar52690	HCV CKS-3	756	7	5.9	36	2	AAW50784	Aaw50784	Peptide u
684	8	6.8	781	4	AAB51372	Ab51372	HCV recom	757	7	5.9	40	2	AAR20772	Aar20772	Peptide v
685	8	6.8	781	4	ABW01858	Abw01858	HCV-CKS-3	758	7	5.9	40	2	AAR74618	Aar74618	Hepatitis
686	8	6.8	1786	1	AAP90158	Ap90158	Protein s	759	7	5.9	60	5	AAW01866	Aaw01866	HCV NS-4
687	8	6.8	3010	9	ADK40793	Adk40793	HCV polym	760	7	5.9	68	5	ABP00036	Abp00036	Human ORF
688	8	6.8	3010	9	ADK40790	Adk40790	HCV polym	761	7	5.9	75	2	AAW22595	Aaw22595	Human hep
689	8	6.8	3013	9	ADK40826	Adk40826	HCV polym	762	7	5.9	87	2	AAR49654	Aar49654	HCV pepti
690	7	5.9	7	2	AAW71274	Aaw71274	Cleavable	763	7	5.9	88	2	AAW44816	Aaw44816	Hepatitis
691	7	5.9	7	6	AAC023378	Aac023378	HCV NS4A	764	7	5.9	93	9	ADZ14570	Adz14570	Human tum
692	7	5.9	8	4	AAJ01353	Aaj01353	Hepatitis	765	7	5.9	107	8	ADY22233	Ady22233	Plant ful
693	7	5.9	8	6	AAO23379	Aao23379	HCV NS4A	766	7	5.9	114	7	ADC96973	Adc96973	E. faeciu
694	7	5.9	9	2	AAW82070	Aaw82070	Fluorogen	767	7	5.9	116	9	ADY52505	Ady52505	Murine Ts
695	7	5.9	9	2	AAJ01788	Aaj01788	Hepatitis	768	7	5.9	125	9	AAR44407	Aar44407	NANBHV de
696	7	5.9	9	4	AAJ03112	Aaj03112	Hepatitis	769	7	5.9	125	2	AAR44406	Aar44406	NANBHV de
697	7	5.9	9	4	AAJ00948	Aaj00948	Hepatitis	770	7	5.9	125	2	AAR44408	Aar44408	NANBHV de
698	7	5.9	9	4	AAJ00436	Aaj00436	Hepatitis	771	7	5.9	125	2	AAR44404	Aar44404	NANBHV de
699	7	5.9	9	4	AAJ03686	Aaj03686	Hepatitis	772	7	5.9	125	2	AAR44405	Aar44405	NANBHV de
700	7	5.9	9	4	AAJ03418	Aaj03418	Hepatitis	773	7	5.9	128	2	AAR63300	Aar63300	Polypepti
701	7	5.9	9	5	ABU60415	Abu60415	Protease	774	7	5.9	130	7	ABO76790	Ab076790	Pseudomon
702	7	5.9	9	6	AAO233380	Aao233380	HCV NS4A	775	7	5.9	131	3	AAO792483	Aao792483	TYLVCV C3
703	7	5.9	9	9	ADZ86594	Adz86594	Cytotoxic f	776	7	5.9	145	7	ADG08302	Adg08302	Novel pro
704	7	5.9	10	2	AAR61536	Aar61536	Peptide f	777	7	5.9	146	4	AAB76733	Aab76733	Corynebac
705	7	5.9	10	2	AAW13785	Aaw13785	Hepatitis	778	7	5.9	147	7	ADM26550	Adm26550	Hyperther
706	7	5.9	10	4	AAJ01927	Aaj01927	Hepatitis	779	7	5.9	153	7	ADE08303	Ad08303	Novel pro
707	7	5.9	10	4	AAJ01067	Aaj01067	Hepatitis	780	7	5.9	161	8	ADX78761	Adx78761	Plant ful
708	7	5.9	10	4	AAJ00662	Aaj00662	Hepatitis	781	7	5.9	163	7	ABO81086	Ab081086	Pseudomon
709	7	5.9	10	6	AAO233381	Aao233381	HCV NS4A	782	7	5.9	163	8	ADX71619	Adx71619	Plant ful
710	7	5.9	10	8	ADE97767	Ad097767	Immunogen	783	7	5.9	171	8	ADY06215	Ady06215	Plant ful
711	7	5.9	11	4	AAJ01079	Aaj01079	Hepatitis	784	7	5.9	172	4	AAG75519	Aag75519	Human col
712	7	5.9	11	4	AAJ02023	Aaj02023	Hepatitis	785	7	5.9	173	4	AAW06677	Aaw06677	Human foe
713	7	5.9	11	4	AAJ00694	Aaj00694	Hepatitis	786	7	5.9	177	8	ADY94276	Ady94276	Plant ful
714	7	5.9	11	6	AAO233382	Aao233382	HCV NS4A	787	7	5.9	177	8	ADY04940	Ady04940	Plant ful
715	7	5.9	13	4	AAB96865	Aab96865	Hepatitis	788	7	5.9	177	8	ADY22585	Ady22585	Plant ful
716	7	5.9	13	4	AAB96863	Aab96863	Hepatitis	789	7	5.9	179	2	AAV41260	Aav41260	Amino aci
717	7	5.9	14	2	AAW09243	Aaw09243	HCV NS4A	790	7	5.9	179	8	ADX78726	Adx78726	Plant ful
718	7	5.9	14	9	ADV23004	Adv23004	HCV H77 i	791	7	5.9	179	8	ADX78414	Adx78414	Plant ful
719	7	5.9	14	9	ADV23002	Adv23002	HCV H77 i	792	7	5.9	180	8	ADX78722	Adx78722	Plant ful
720	7	5.9	14	9	ADV22999	Adv22999	HCV H77 i	793	7	5.9	180	8	ADX78423	Adx78423	Plant ful
721	7	5.9	15	9	ADV23003	Adv23003	HCV H77 i	794	7	5.9	183	9	ABM92257	Abm92257	M. xanthu
722	7	5.9	17	5	AAU90051	Aau90051	Insulin/i	795	7	5.9	188	9	ABM90870	Abm90870	M. xanthu
723	7	5.9	18	2	AAW37940	Aaw37940	HCV NS-4	796	7	5.9	201	2	AAV41261	Aav41261	Amino aci
724	7	5.9	18	2	AAW82171	Aaw82171	Fluorogen	797	7	5.9	202	8	ADX33690	Adx33690	Plant ful
725	7	5.9	18	4	AAW73170	Aaw73170	Protease	798	7	5.9	209	8	ADK46560	Adk46560	Streptoco
726	7	5.9	18	4	AAW73173	Aaw73173	Protease	799	7	5.9	210	4	ABB67469	Abb67469	Drosophi
727	7	5.9	18	8	ADN88410	Adn88410	Fluorogen	800	7	5.9	217	3	AAV75063	Aav75063	Neisseria
728	7	5.9	18	9	ADV23007	Adv23007	HCV H77 i	801	7	5.9	217	3	AAV75064	Aav75064	Neisseria
729	7	5.9	18	9	ADV23015	Adv23015	HCV H77 i	802	7	5.9	217	3	AAV75062	Aav75062	Neisseria
730	7	5.9	18	9	ADV23014	Adv23014	HCV H77 i	803	7	5.9	217	5	AAU72926	Aau72926	Neisseria
731	7	5.9	19	2	AAW30982	Aaw30982	Short HCV	804	7	5.9	217	6	ABP77518	Abp77518	N. gonorr
732	7	5.9	19	2	AAW82174	Aaw82174	Fluorogen	805	7	5.9	222	2	AAW50078	Aaw50078	NANBH vir
733	7	5.9	19	4	AAW73176	Aaw73176	Protease	806	7	5.9	222	2	AAW49655	Aaw49655	HCV pepti
734	7	5.9	19	8	ADN88413	Adn88413	Fluorogen	807	7	5.9	222	4	AAR49655	Aar49655	Coryneb
735	7	5.9	20	2	AAW74610	Aaw74610	Hepatitis	808	7	5.9	222	4	AAG90702	Aag90702	C glutami
736	7	5.9	20	2	AAW82177	Aaw82177	Fluorogen	809	7	5.9	222	6	ABU02260	Abu02260	S. pneumo
737	7	5.9	20	8	ADN88416	Adn88416	Fluorogen	810	7	5.9	222	6	ABU46216	Abu46216	Protein e
738	7	5.9	22	2	AAW41149	Aaw41149	HCV (type	811	7	5.9	222	8	ADJ81683	Adj81683	Non-A-non
739	7	5.9	22	2	AAW41150	Aaw41150	HCV (type	812	7	5.9	225	7	ADR95020	Adr95020	Novel S.
740	7	5.9	22	2	AAW99902	Aaw99902	Hepatitis	813	7	5.9	227	9	AEA58890	Aea58890	Streptoco
741	7	5.9	22	2	AAW99901	Aaw99901	Hepatitis	814	7	5.9	230	8	ADJ77181	Adj77181	Plant ful
742	7	5.9	22	8	ADN00916	Adn00916	Hepatitis	815	7	5.9	245	8	ADJ79593	Adj79593	Hepatitis
743	7	5.9	22	8	ADN00917	Adn00917	Hepatitis	816	7	5.9	250	7	ADC38707	Adc38707	Human sec
744	7	5.9	25	2	AAW37388	Aaw37388	Peptide u	817	7	5.9	255	6	ABU22837	Abu22837	Protein e
745	7	5.9	25	2	AAW50788	Aaw50788	Hepatitis	818	7	5.9	255	7	ADA49405	Ada49405	Multi-epi
746	7	5.9	26	2	AAW16440	Aaw16440	Human gro	819	7	5.9	255	8	ADZ40483	Adz40483	Epigene c
747	7	5.9	26	2	AAW52785	Aaw52785	Human gro	820	7	5.9	255	9	ADZ40585	Adz40585	HCV.1 mul
748	7	5.9	27	2	AAW16439	Aaw16439	Human gro	821	7	5.9	256	2	AAW20611	Aaw20611	C10-13 NA
749	7	5.9	27	2	AAW52784	Aaw52784	Human gro	822	7	5.9	260	3	AAW43436	Aaw43436	Human can
750	7	5.9	28	2	AAW16419	Aaw16419	Human gro	823	7	5.9	260	4	AAU64862	Aau64862	Propionib
751	7	5.9	28	2	AAW52764	Aaw52764	Human gro	824	7	5.9	260	5	AAE19895	Aae19895	Hepatitis
752	7	5.9	30	5	AAU84714	Aau84714	HCV HepC1	825	7	5.9	260	6	ABM61381	Abm61381	Propionib
753	7	5.9	34	2	AAV16755	Aav16755	Calcitoni	826	7	5.9	260	7	ABW00346	Abw00346	Hepatitis
754	7	5.9	35	3	AAV67801	Aav67801	Peptide #	827	7	5.9	262	6	ABU43166	Abu43166	Protein e

828	7	5.9	262	6	ABU19953	Abu19953 Protein e	901	7	5.9	605	8	ADN19948	Adn19948 Bacterial
829	7	5.9	263	5	ADP39361	Adp39361 Staphyloc	902	7	5.9	607	8	ADS29014	Ads29014 Bacterial
830	7	5.9	263	8	ADS04531	Ads04531 Staphyloc	903	7	5.9	607	8	ADS21496	Ads21496 Bacterial
831	7	5.9	266	8	ADU02681	Adu02681 Novel hum	904	7	5.9	622	5	ABP64821	Abp64821 Human pro
832	7	5.9	269	2	AAW92815	Aaw92815 HCV NS4B	905	7	5.9	622	5	ADQ95978	Adq95978 T cell ac
833	7	5.9	273	3	AAG27984	Aag27984 Arabidops	906	7	5.9	626	5	AAU74619	Aau74619 Oestrogen
834	7	5.9	273	8	ADR04184	Adr04184 E faecium	907	7	5.9	626	5	ABG96322	Abg96322 Human ova
835	7	5.9	277	2	AAR54203	Aar54203 snab gene	908	7	5.9	626	5	ABP70108	Abp70108 Human NOV
836	7	5.9	277	3	AAG27983	Aag27983 Arabidops	909	7	5.9	626	8	ADU06701	Adu06701 Novel bro
837	7	5.9	277	6	ABU279809	Abu279809 Protein e	910	7	5.9	647	4	AAE06574	Aae06574 Human pro
838	7	5.9	288	8	ADX96062	Adx96062 Plant ful	911	7	5.9	647	4	AAE06574	Aae06574 Human pro
839	7	5.9	294	7	ADE08670	Ade08670 Novel pro	912	7	5.9	647	4	ABW74710	Abw74710 Human mem
840	7	5.9	295	2	RAR90027	Rar90027 Methionin	913	7	5.9	647	4	AAE01677	Aae01677 Human gen
841	7	5.9	295	8	ADN46674	Adn46674 Thermococ	914	7	5.9	647	5	AAU74618	Aau74618 Oestrogen
842	7	5.9	306	4	ADG81135	Adg81135 Mycobacte	915	7	5.9	647	5	ABG63939	Abg63939 Human alb
843	7	5.9	316	4	ABB68498	Abb68498 Drosophil	916	7	5.9	647	7	ADE31719	Ade31719 Human 646
844	7	5.9	319	3	AY92506	Aay92506 Human OXR	917	7	5.9	647	7	ADN40019	Adn40019 Cancer/an
845	7	5.9	319	4	AAB94861	Aab94861 Human pro	918	7	5.9	647	8	ADL77204	Adl77204 Albumin f
846	7	5.9	321	3	AAG27982	Aag27982 Arabidops	919	7	5.9	647	8	ADQ95976	Adq95976 T cell ac
847	7	5.9	324	7	ADE08132	Ade08132 Novel pro	920	7	5.9	647	8	ADR46689	Adr46689 Cancer-as
848	7	5.9	335	8	ADN26699	Adn26699 Bacterial	921	7	5.9	647	9	ADZ51373	Adz51373 Amino aci
849	7	5.9	338	6	ABM65633	Abm65633 Propionib	922	7	5.9	647	9	ABE42807	Abe42807 Snail tra
850	7	5.9	341	8	ADN26623	Adn26623 Bacterial	923	7	5.9	665	6	ABU49815	Abu49815 Protein e
851	7	5.9	344	5	ABH89956	Abh89956 Human pol	924	7	5.9	673	4	ABG26099	Abg26099 Novel hum
852	7	5.9	351	4	ADG64355	Adg64355 Human lam	925	7	5.9	707	6	ABU33621	Abu33621 Protein e
853	7	5.9	357	8	ADU02842	Adu02842 Novel hum	926	7	5.9	707	9	ABE38126	Abe38126 L. pneumo
854	7	5.9	359	8	ADS28712	Ads28712 Bacterial	927	7	5.9	707	9	ABE41417	Abe41417 L. pneumo
855	7	5.9	361	6	ABU23531	Abu23531 Protein e	928	7	5.9	717	4	ABG06939	Abg06939 Novel hum
856	7	5.9	363	4	AAE09651	Aae09651 Human gen	929	7	5.9	737	7	ADC87583	Adc87583 Human GPC
857	7	5.9	363	7	ABW01087	Abw01087 Human gen	930	7	5.9	740	6	ABU48461	Abu48461 Protein e
858	7	5.9	366	2	AAI40499	Aay40499 Synchocy	931	7	5.9	774	8	ADN25720	Adn25720 Bacterial
859	7	5.9	367	3	AAI13718	Aab19718 Arabidops	932	7	5.9	785	4	AAG91691	Aag91691 C glutami
860	7	5.9	367	7	ADC35637	Adc35637 Arabidops	933	7	5.9	808	7	ABO83221	Abog83221 Pseudomon
861	7	5.9	379	4	ABG29922	Abg29922 Novel hum	934	7	5.9	816	2	AAR44143	Aar44143 Rabbit so
862	7	5.9	389	9	ABE14376	Abel14376 Plant lip	935	7	5.9	816	2	ABM84568	Abm84568 Human dia
863	7	5.9	396	2	AAW37377	Aaw37377 Hepatitis	936	7	5.9	817	8	ADN26407	Adn26407 Bacterial
864	7	5.9	396	3	AAI80193	Aay80193 Hepatitis	937	7	5.9	870	6	ADB99968	Adb99968 Enterohae
865	7	5.9	409	7	ABO81715	Abog81715 Pseudomon	938	7	5.9	879	6	ADBI2249	Adbi2249 Alloiococ
866	7	5.9	411	6	ABP80861	Abp80861 N. gonorr	939	7	5.9	881	7	ADC00520	Adc00520 Enterohae
867	7	5.9	414	6	ABU49671	Abu49671 Protein e	940	7	5.9	913	7	ABO80616	Abog80616 Pseudomon
868	7	5.9	418	8	ADS28776	Ads28776 Bacterial	941	7	5.9	1036	7	ADB64658	Adb64658 Human pro
869	7	5.9	421	2	AAW37376	Aaw37376 Hepatitis	942	7	5.9	1082	6	ABU21740	Abu21740 Protein e
870	7	5.9	421	6	ABP78409	Abp78409 N. gonorr	943	7	5.9	1116	7	ADF04094	Adf04094 Bacterial
871	7	5.9	421	7	ADI21224	Adi21224 Novel hum	944	7	5.9	1255	7	ADJ87387	Adj87387 DNA repli
872	7	5.9	426	4	AAU27812	Aau27812 Human ful	945	7	5.9	1255	7	ADL65837	Adl65837 C. glutam
873	7	5.9	426	4	AAU42357	Aau42357 Propionib	946	7	5.9	2010	5	AAU84801	Aau84801 HCV HepC
874	7	5.9	426	6	ABM38876	Abm38876 Propionib	947	7	5.9	2278	3	AAI53677	Aay53677 Sequence
875	7	5.9	428	6	ABU28800	Abu28800 Protein e	948	7	5.9	2408	6	ABU70437	Abu70437 Human adi
876	7	5.9	431	6	ABU37936	Abu37936 Protein e	949	7	5.9	2458	4	AAI39071	Aai39071 Human pol
877	7	5.9	431	8	ADP08253	Adp08253 Neisseria	950	7	5.9	2458	4	AAU28088	Aau28088 Novel hum
878	7	5.9	433	6	ADA36359	Ada36359 Acinetoba	951	7	5.9	2940	7	ADC42917	Adc42917 Hepatitis
879	7	5.9	438	7	ABO61746	Abog61746 Klebsiell	952	7	5.9	3006	7	ADE40130	Ade40130 Human NOV
880	7	5.9	441	8	ADS03030	Ads03030 Bacterial	953	7	5.9	3033	2	AAR33538	Aar33538 NANBH vir
881	7	5.9	452	8	ADG42560	Adg42560 Bacterial	954	7	5.9	3033	4	AAI59172	Aai59172 Protein e
882	7	5.9	453	6	ABU49575	Abu49575 Protein e	955	7	5.9	3033	4	AAI31168	Aai31168 Amino aci
883	7	5.9	457	6	ABU22514	Abu22514 Protein e	956	7	5.9	3033	5	ABG30688	Abg30688 Human HCV
884	7	5.9	457	7	ABO73936	Abog73936 Pseudomon	957	7	5.9	3033	9	ADV04740	Adv04740 Hepatitis
885	7	5.9	458	8	ADX96169	Adx96169 Plant ful	958	7	5.9	3033	9	ADV04738	Adv04738 Hepatitis
886	7	5.9	473	7	ADJ53182	Adj53182 Rebeccamy	959	7	5.9	3033	9	ADX40815	Adx40815 HCV polym
887	7	5.9	477	9	ABM92710	Abm92710 M. xanthu	960	7	5.9	3033	9	ADX40819	Adx40819 HCV polym
888	7	5.9	493	8	ADT89472	Adt89472 Caenorhab	961	7	5.9	3037	9	ADX40778	Adx40778 HCV polym
889	7	5.9	524	8	ADP04515	Adp04515 Sea squir	962	7	5.9	3262	3	AAI53675	Aay53675 Mechanica
890	7	5.9	526	8	ADU48513	Adu48513 Meloidogy	963	7	5.9	3264	3	AAI53676	Aay53676 Protein 2
891	7	5.9	526	8	ADU48511	Adu48511 Meloidogy	964	7	5.9	5183	6	AAI14793	Aai14793 Human mic
892	7	5.9	530	6	ABM67618	Abm67618 Photorhab	965	7	5.9	5183	8	ADO44006	Ado44006 Amino aci
893	7	5.9	536	4	AAU27984	Aau27984 Human con	966	7	5.9	5183	9	ADX06154	Adx06154 Cyclin-de
894	7	5.9	550	4	AAU27984	Aau27984 Human con	967	7	5.9	5644	6	AAE36120	Aae36120 Streptomy
895	7	5.9	553	9	ADZ51375	Adz51375 Amino aci	968	7	5.9	7349	6	ABU11385	Abu11385 Protein e
896	7	5.9	563	3	ABM10080	Abm10080 F. bident	969	6	5.1	6	2	AAI28715	Aay28715 Human hep
897	7	5.9	564	3	ABM10081	Abm10081 S. olerac	970	6	5.1	6	2	AAI28714	Aay28714 Human hep
898	7	5.9	567	4	ABB65162	Abb65162 Drosophil	971	6	5.1	6	2	AAI28740	Aay28740 Human hep
899	7	5.9	585	8	ADS28277	Ads28277 Bacterial	972	6	5.1	6	2	AAI28754	Aay28754 Human hep
900	7	5.9	594	8	ADH12819	Adh12819 Abalone (973	6	5.1	6	3	AAI97454	Aay97454 Hepatitis

974	6	5.1	6	3	AAJ97437	Aay97437 Hepatitis	1047	6	5.1	10	4	AAJ01933	Aaj01933 Hepatitis
975	6	5.1	6	3	AAJ97436	Aay97436 Hepatitis	1048	6	5.1	10	4	AAJ00567	Aaj00567 Hepatitis
976	6	5.1	6	3	AAJ023377	Aad23377 HCV NS4A	1049	6	5.1	10	4	AAJ00273	Aaj00273 Hepatitis
977	6	5.1	6	8	ADU39858	Adu39858 Peptide i	1050	6	5.1	10	4	AAJ00693	Aaj00693 Hepatitis
978	6	5.1	6	9	ADV78811	Adv78811 HCV inhib	1051	6	5.1	10	4	AAJ02603	Aaj02603 Hepatitis
979	6	5.1	6	9	ADW97223	Adw97223 HCV NS3 p	1052	6	5.1	10	4	AAJ01865	Aaj01865 Hepatitis
980	6	5.1	6	9	ABE25213	Aeb25213 Peptide i	1053	6	5.1	10	5	ABB95155	Abb95155 CTL epit
981	6	5.1	6	2	AAJ35974	Aar35974 Hepatitis	1054	6	5.1	10	5	ABB94858	Abb94858 CTL epit
982	6	5.1	6	4	AAJ01075	Aaj01075 Hepatitis	1055	6	5.1	10	5	ABB94971	Abb94971 CTL epit
983	6	5.1	6	4	AAJ00435	Aaj00435 Hepatitis	1056	6	5.1	10	5	ABU60413	Abu60413 Protease
984	6	5.1	6	4	AAJ00683	Aaj00683 Hepatitis	1057	6	5.1	10	7	ADA49567	Ada49567 Multi-epi
985	6	5.1	6	8	AAJ01727	Aaj01727 Hepatitis	1058	6	5.1	10	7	ADe79969	AdE79969 Plasmodiu
986	6	5.1	6	4	AAJ01667	Aaj01667 Hepatitis	1059	6	5.1	10	7	ADL18160	Adl18160 Hepatitis
987	6	5.1	6	4	AAJ00675	Aaj00675 Hepatitis	1060	6	5.1	10	7	ADW32963	Adw32963 HLA bindi
988	6	5.1	6	4	AAJ00566	Aaj00566 Hepatitis	1061	6	5.1	10	8	ADL71857	Adl71857 Hepatitis
989	6	5.1	6	4	AAJ01377	Aaj01377 Hepatitis	1062	6	5.1	10	8	ADO24245	Ado24245 P falcipa
990	6	5.1	6	4	AAJ02350	Aaj02350 Hepatitis	1063	6	5.1	10	8	ADN02814	Adn02814 Hepatitis
991	6	5.1	6	2	AAW49391	Aaw49391 Human leu	1064	6	5.1	10	8	ADN02818	Adn02818 Hepatitis
992	6	5.1	6	2	AAW43337	Aaw43337 Immunogen	1065	6	5.1	10	9	ADY53188	Ady53188 Hepatitis
993	6	5.1	6	2	AAJ47503	Aay47503 Immunogen	1066	6	5.1	10	9	ADY84772	Ady84772 Hepatitis
994	6	5.1	6	3	AAJ73108	Aay73108 Hepatitis	1067	6	5.1	10	9	ADZ40747	Adz40747 Multi-epi
995	6	5.1	6	3	AAJ73304	Aay73304 Plasmodiu	1068	6	5.1	11	2	AAW13793	Aaw13793 Cleavable
996	6	5.1	6	4	AAU26982	Aau26982 Human leu	1069	6	5.1	11	2	AAW171277	Aaw171277
997	6	5.1	6	4	AAU26649	Aau26649 Human leu	1070	6	5.1	11	4	AAJ01597	Aaj01597 Hepatitis
998	6	5.1	6	4	AAJ02575	Aaj02575 Hepatitis	1071	6	5.1	11	4	AAJ01945	Aaj01945 Hepatitis
999	6	5.1	6	4	AAJ01832	Aaj01832 Hepatitis	1072	6	5.1	11	4	AAJ00676	Aaj00676 Hepatitis
1000	6	5.1	6	4	AAJ03208	Aaj03208 Hepatitis	1073	6	5.1	11	4	AAJ00862	Aaj00862 Hepatitis
1001	6	5.1	6	4	AAJ00034	Aaj00034 Hepatitis	1074	6	5.1	11	4	AAJ00225	Aaj00225 Hepatitis
1002	6	5.1	6	4	AAJ04088	Aaj04088 Hepatitis	1075	6	5.1	11	4	AAJ01584	Aaj01584 Hepatitis
1003	6	5.1	6	4	AAJ03212	Aaj03212 Hepatitis	1076	6	5.1	11	4	AAJ01590	Aaj01590 Hepatitis
1004	6	5.1	6	4	AAJ01836	Aaj01836 Hepatitis	1077	6	5.1	11	7	ADW37382	Adw37382 HLA bindi
1005	6	5.1	6	4	AAJ03515	Aaj03515 Hepatitis	1078	6	5.1	11	7	ADW37378	Adw37378 HLA bindi
1006	6	5.1	6	4	AAJ03519	Aaj03519 Hepatitis	1079	6	5.1	11	7	ADW31336	Adw31336 HLA bindi
1007	6	5.1	6	4	AAJ03868	Aaj03868 Hepatitis	1080	6	5.1	11	7	ADW32964	Adw32964 HLA bindi
1008	6	5.1	6	4	AAJ00272	Aaj00272 Hepatitis	1081	6	5.1	12	2	AAW13777	Aaw13777 Hepatitis
1009	6	5.1	6	4	AAJ03604	Aaj03604 Hepatitis	1082	6	5.1	12	2	AAW13776	Aaw13776 Hepatitis
1010	6	5.1	6	4	AAJ00876	Aaj00876 Hepatitis	1083	6	5.1	12	2	AAW47138	Aaw47138 Hepatitis
1011	6	5.1	6	4	AAJ03600	Aaj03600 Hepatitis	1084	6	5.1	12	2	AAW94437	Aaw94437 Mutant pr
1012	6	5.1	6	4	AAJ00661	Aaj00661 Hepatitis	1085	6	5.1	12	3	AAE57197	Aay57197 HCV NS3 p
1013	6	5.1	6	4	AAJ00684	Aaj00684 Hepatitis	1086	6	5.1	12	5	AAE21850	Aae21850 Hepatitis
1014	6	5.1	6	4	AAJ03789	Aaj03789 Hepatitis	1087	6	5.1	12	5	AAE19904	Aae19904 Hepatitis
1015	6	5.1	6	4	AAJ01754	Aaj01754 Hepatitis	1088	6	5.1	12	6	AAO23352	Aao23352 Hepatitis
1016	6	5.1	6	4	AAJ03966	Aaj03966 Hepatitis	1089	6	5.1	12	6	ABW00355	Abw00355 Hepatitis
1017	6	5.1	6	5	ABB95096	Abb95096 CTL epit	1090	6	5.1	12	8	ADG47672	Adg47672 HCV NS3/4
1018	6	5.1	6	5	ABB94696	Abb94696 CTL epit	1091	6	5.1	12	8	ADM01195	Adm01195 Targeted
1019	6	5.1	6	5	ABB94774	Abb94774 CTL epit	1092	6	5.1	12	8	ADT88719	Adt88719 Recombina
1020	6	5.1	6	5	AAU95866	Aau95866 Immunogen	1093	6	5.1	13	2	AAW09244	Aaw09244 HCV NS4A
1021	6	5.1	6	5	ADA49530	Ada49530 Multi-epi	1094	6	5.1	13	2	AAW85699	Aaw85699 HCV-J NS4
1022	6	5.1	6	7	ADM32537	Adw32537 HLA bindi	1095	6	5.1	13	2	AAW13775	Aaw13775 Hepatitis
1023	6	5.1	6	8	ADe97626	AdE97626 Immunogen	1096	6	5.1	13	2	AAW13786	Aaw13786 Hepatitis
1024	6	5.1	6	9	ADJ79565	Adj79565 CTL epit	1097	6	5.1	13	2	AAW47145	Aaw47145 Hepatitis
1025	6	5.1	6	9	ADK04446	Adk04446 Hepatitis	1098	6	5.1	13	2	AAW47154	Aaw47154 Hepatitis
1026	6	5.1	6	9	ADO24208	Ado24208 HCV epit	1099	6	5.1	13	2	AAW37809	Aaw37809 Peptide d
1027	6	5.1	6	8	ADU99948	Adu99948 BfY3 tumo	1100	6	5.1	13	2	AAV28783	Aay28783 Hepatitis
1028	6	5.1	6	9	ADV00007	Adv00007 BfY3 tumo	1101	6	5.1	13	3	AAV57213	Aay57213 HCV NS3 p
1029	6	5.1	6	9	ADU99949	Adu99949 BfY3 tumo	1102	6	5.1	13	3	AAV57204	Aay57204 HCV NS4A
1030	6	5.1	6	9	ADU99947	Adu99947 BfY3 tumo	1103	6	5.1	13	3	AAV54449	Aay54449 Substrate
1031	6	5.1	6	9	ADX08730	Adx08730 Hepatitis	1104	6	5.1	13	3	AAV99553	Aay99553 Hepatitis
1032	6	5.1	6	9	ADZ40710	Adz40710 Multi-epi	1105	6	5.1	13	4	AAV97115	Aab97115 Hepatitis
1033	6	5.1	10	2	AAW84605	Aar84605 HCV-1 der	1106	6	5.1	13	5	ABB05364	Abb05364 Substrate
1034	6	5.1	10	2	AAW13782	Aaw13782 Hepatitis	1107	6	5.1	14	2	AAW32962	Aar32962 Mastopara
1035	6	5.1	10	2	AAW13780	Aaw13780 Hepatitis	1108	6	5.1	14	2	AAW32954	Aar32954 Mastopara
1036	6	5.1	10	2	AAW13783	Aaw13783 Hepatitis	1109	6	5.1	14	2	AAW32963	Aar32963 Mastopara
1037	6	5.1	10	2	AAW13781	Aaw13781 Hepatitis	1110	6	5.1	14	2	AAW32944	Aar32944 Mastopara
1038	6	5.1	10	2	AAW13779	Aaw13779 Hepatitis	1111	6	5.1	14	2	AAW13774	Aaw13774 Hepatitis
1039	6	5.1	10	2	AAW13778	Aaw13778 Hepatitis	1112	6	5.1	14	7	ADG73516	Adg73516 E faecali
1040	6	5.1	10	2	AAW14711	Aaw14711 Hepatitis	1113	6	5.1	14	9	AEC11116	Aec11116 Enterococ
1041	6	5.1	10	2	AAW82069	Aaw82069 Fluorogen	1114	6	5.1	15	2	AAW13773	Aaw13773 Hepatitis
1042	6	5.1	10	2	AAV47513	Aay47513 Immunogen	1115	6	5.1	15	2	AAW85168	Aaw85168 Helper T-
1043	6	5.1	10	3	AAU73289	Aay73289 Plasmodiu	1116	6	5.1	15	2	AAW85179	Aaw85179 Helper T-
1044	6	5.1	10	4	AAU26653	Aau26653 Human leu	1117	6	5.1	15	2	AAW85374	Aaw85374 Helper T-
1045	6	5.1	10	4	AAU26986	Aau26986 Human leu	1118	6	5.1	15	2	AAW85153	Aaw85153 Helper T-
1046	6	5.1	10	4	AAJ01526	Aaj01526 Hepatitis	1119	6	5.1	15	2	AAW85431	Aaw85431 Helper T-

1120	6	5.1	15	3	AAV88514	Aay88514 Peptide #	1193	6	5.1	26	2	AA14913	Aar14913 Antigen j
1121	6	5.1	15	3	AAV73279	Aay73279 Plasmid	1194	6	5.1	26	2	AA14912	Aar14912 Antigen j
1122	6	5.1	15	3	AAV703257	Aaj03257 Hepatitis	1195	6	5.1	26	3	AA54383	Aay54383 Amino aci
1123	6	5.1	15	4	AAJ03380	Aaj03380 Hepatitis	1196	6	5.1	26	5	ABG95818	Abg95818 Cell pene
1124	6	5.1	15	4	AAJ03074	Aaj03074 Hepatitis	1197	6	5.1	26	6	AAO16886	Aao16886 Cell-perm
1125	6	5.1	15	4	AAJ033562	Aaj033562 Hepatitis	1198	6	5.1	27	3	AAO16886	Aao16886 Cell-perm
1126	6	5.1	15	4	AAJ033647	Aaj033647 Hepatitis	1199	6	5.1	27	3	AA54384	Aay54384 Amino aci
1127	6	5.1	15	4	AAJ04007	Aaj04007 Hepatitis	1200	6	5.1	27	3	AA54392	Aay54392 Amino aci
1128	6	5.1	15	7	ADA49597	Ada49597 Multi-epi	1201	6	5.1	28	3	AA54401	Aay54401 Amino aci
1129	6	5.1	15	7	ADG73517	Adg73517 E faecal	1202	6	5.1	28	3	AA54385	Aay54385 Amino aci
1130	6	5.1	15	7	ADW36071	Adw36071 HLA bindi	1203	6	5.1	28	3	AA54393	Aay54393 Amino aci
1131	6	5.1	15	7	ADW36553	Adw36553 HLA bindi	1204	6	5.1	28	8	ADK16011	Adk16011 African c
1132	6	5.1	15	7	ADW35185	Adw35185 HLA bindi	1205	6	5.1	28	8	ADO81244	Ado81244 Protein s
1133	6	5.1	15	7	ADW35959	Adw35959 HLA bindi	1206	6	5.1	29	2	AA22287	Aar22287 Hepatitis
1134	6	5.1	15	7	ADW35902	Adw35902 HLA bindi	1207	6	5.1	29	2	AA67298	Aay67298 Hepatitis
1135	6	5.1	15	7	ADW36128	Adw36128 HLA bindi	1208	6	5.1	29	3	AA54394	Aay54394 Amino aci
1136	6	5.1	15	7	ADW36543	Adw36543 HLA bindi	1209	6	5.1	29	3	AA54402	Aay54402 Amino aci
1137	6	5.1	15	7	ADW33952	Adw33952 HLA bindi	1210	6	5.1	29	3	AA54386	Aay54386 Amino aci
1138	6	5.1	15	7	ADW36551	Adw36551 HLA bindi	1211	6	5.1	30	2	AA67299	Aay67299 Hepatitis
1139	6	5.1	15	8	ADL26392	Adl26392 Synthetic	1212	6	5.1	30	3	AA54395	Aay54395 Amino aci
1140	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1213	6	5.1	30	3	AA54403	Aay54403 Amino aci
1141	6	5.1	15	8	ADL26390	Adl26390 Synthetic	1214	6	5.1	30	3	AA54387	Aay54387 Amino aci
1142	6	5.1	15	8	ADL26342	Adl26342 Synthetic	1215	6	5.1	30	5	AAU84710	Aau84710 HCV HepC
1143	6	5.1	15	8	ADL26388	Adl26388 Synthetic	1216	6	5.1	30	9	ADV99750	Adv99750 Glucanase
1144	6	5.1	15	8	ADL26389	Adl26389 Synthetic	1217	6	5.1	31	3	AA54388	Aay54388 Amino aci
1145	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1218	6	5.1	31	3	AA54404	Aay54404 Amino aci
1146	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1219	6	5.1	31	3	AA54396	Aay54396 Amino aci
1147	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1220	6	5.1	31	4	AAO08727	Aao08727 Human pol
1148	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1221	6	5.1	31	4	AAO08727	Aao08727 Human pol
1149	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1222	6	5.1	31	4	AAO08727	Aao08727 Human pol
1150	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1223	6	5.1	31	4	AAO08727	Aao08727 Human pol
1151	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1224	6	5.1	31	4	AAO08727	Aao08727 Human pol
1152	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1225	6	5.1	31	4	AAO08727	Aao08727 Human pol
1153	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1226	6	5.1	31	4	AAO08727	Aao08727 Human pol
1154	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1227	6	5.1	31	4	AAO08727	Aao08727 Human pol
1155	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1228	6	5.1	31	4	AAO08727	Aao08727 Human pol
1156	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1229	6	5.1	31	4	AAO08727	Aao08727 Human pol
1157	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1230	6	5.1	31	4	AAO08727	Aao08727 Human pol
1158	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1231	6	5.1	31	4	AAO08727	Aao08727 Human pol
1159	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1232	6	5.1	31	4	AAO08727	Aao08727 Human pol
1160	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1233	6	5.1	31	4	AAO08727	Aao08727 Human pol
1161	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1234	6	5.1	31	4	AAO08727	Aao08727 Human pol
1162	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1235	6	5.1	31	4	AAO08727	Aao08727 Human pol
1163	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1236	6	5.1	31	4	AAO08727	Aao08727 Human pol
1164	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1237	6	5.1	31	4	AAO08727	Aao08727 Human pol
1165	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1238	6	5.1	31	4	AAO08727	Aao08727 Human pol
1166	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1239	6	5.1	31	4	AAO08727	Aao08727 Human pol
1167	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1240	6	5.1	31	4	AAO08727	Aao08727 Human pol
1168	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1241	6	5.1	31	4	AAO08727	Aao08727 Human pol
1169	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1242	6	5.1	31	4	AAO08727	Aao08727 Human pol
1170	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1243	6	5.1	31	4	AAO08727	Aao08727 Human pol
1171	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1244	6	5.1	31	4	AAO08727	Aao08727 Human pol
1172	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1245	6	5.1	31	4	AAO08727	Aao08727 Human pol
1173	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1246	6	5.1	31	4	AAO08727	Aao08727 Human pol
1174	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1247	6	5.1	31	4	AAO08727	Aao08727 Human pol
1175	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1248	6	5.1	31	4	AAO08727	Aao08727 Human pol
1176	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1249	6	5.1	31	4	AAO08727	Aao08727 Human pol
1177	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1250	6	5.1	31	4	AAO08727	Aao08727 Human pol
1178	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1251	6	5.1	31	4	AAO08727	Aao08727 Human pol
1179	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1252	6	5.1	31	4	AAO08727	Aao08727 Human pol
1180	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1253	6	5.1	31	4	AAO08727	Aao08727 Human pol
1181	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1254	6	5.1	31	4	AAO08727	Aao08727 Human pol
1182	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1255	6	5.1	31	4	AAO08727	Aao08727 Human pol
1183	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1256	6	5.1	31	4	AAO08727	Aao08727 Human pol
1184	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1257	6	5.1	31	4	AAO08727	Aao08727 Human pol
1185	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1258	6	5.1	31	4	AAO08727	Aao08727 Human pol
1186	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1259	6	5.1	31	4	AAO08727	Aao08727 Human pol
1187	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1260	6	5.1	31	4	AAO08727	Aao08727 Human pol
1188	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1261	6	5.1	31	4	AAO08727	Aao08727 Human pol
1189	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1262	6	5.1	31	4	AAO08727	Aao08727 Human pol
1190	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1263	6	5.1	31	4	AAO08727	Aao08727 Human pol
1191	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1264	6	5.1	31	4	AAO08727	Aao08727 Human pol
1192	6	5.1	15	8	ADL26391	Adl26391 Synthetic	1265	6	5.1	31	4	AAO08727	Aao08727 Human pol

1266	6	5.1	62	2	AAV12367	Aay12367 Human 5'	1339	6	5.1	100	4	AA19979	Ab19979 Rous sarc
1267	6	5.1	63	2	AAR49653	Aar49653 HCV pepti	1340	6	5.1	100	8	ADF29125	Adf29125 Thale cre
1268	6	5.1	63	2	AAR49651	Aar49651 HCV pepti	1341	6	5.1	101	3	AG02777	Ag02777 Human sec
1269	6	5.1	63	4	AAU57538	Aau57538 Propionib	1342	6	5.1	101	6	ABP57632	Abp57632 S. muraya
1270	6	5.1	63	6	ABM54057	Abm54057 Propionib	1343	6	5.1	102	7	ADF18733	Adf18733 DNA-depen
1271	6	5.1	64	4	AAU58646	Aau58646 Propionib	1344	6	5.1	103	5	ABP08771	Abp08771 Human ORF
1272	6	5.1	64	6	ABM55165	Abm55165 Propionib	1345	6	5.1	104	4	ABG08601	Abg08601 Novel hum
1273	6	5.1	66	7	ADF58745	Adf58745 Human pol	1346	6	5.1	105	2	AAR2088	Aar2088 Src SH2 r
1274	6	5.1	68	4	AAU65265	Aau65265 Propionib	1347	6	5.1	105	7	ADC88803	Adc88803 Ribosomal
1275	6	5.1	68	4	ABG07328	Abg07328 Novel hum	1348	6	5.1	105	7	ADC88118	Adc88118 Ribosomal
1276	6	5.1	68	6	ADA57464	Ada57464 Human sec	1349	6	5.1	106	6	ADA57463	Ada57463 Human sec
1277	6	5.1	68	6	ABM61784	Abm61784 Propionib	1350	6	5.1	106	6	ADA41343	Ada41343 Human sec
1278	6	5.1	68	6	ADA41344	Ada41344 Human sec	1351	6	5.1	106	6	ABR48065	Abra48065 Human sec
1279	6	5.1	68	6	ABR48066	Abra48066 Human sec	1352	6	5.1	106	7	ADA49411	Ada49411 Multi-epi
1280	6	5.1	69	2	AAR12757	Aar12757 Part of N	1353	6	5.1	106	8	ADO24089	Ado24089 Epigene c
1281	6	5.1	69	2	AAR12758	Aar12758 Part of N	1354	6	5.1	106	9	ADM38949	Adm38949 Novel pla
1282	6	5.1	69	4	AAU34850	Aau34850 E. coli c	1355	6	5.1	106	9	ADZ40591	Adz40591 HCV.3a2 m
1283	6	5.1	69	6	ABU28859	Abu28859 Protein e	1356	6	5.1	107	4	AAU56379	Aau56379 Propionib
1284	6	5.1	69	8	ABO57094	Ab057094 Human gen	1357	6	5.1	107	6	ABM52898	Abm52898 Propionib
1285	6	5.1	70	3	AG02607	Ag02607 Human sec	1358	6	5.1	107	9	ADZ22236	Adz22236 Full leng
1286	6	5.1	70	4	AAO10516	Aao10516 Human pol	1359	6	5.1	108	2	AAU04872	Aau04872 Mycobacte
1287	6	5.1	70	4	AAU59619	Aau59619 Propionib	1360	6	5.1	108	4	ABG20561	Abg20561 Novel hum
1288	6	5.1	70	6	ABM56138	Abm56138 Propionib	1361	6	5.1	109	7	ADF60472	Adf60472 Human con
1289	6	5.1	72	3	AAG03674	Ag03674 Human sec	1362	6	5.1	110	2	AAW38477	Aaw38477 Streptoco
1290	6	5.1	73	4	AAO07892	Aao07892 Human pol	1363	6	5.1	110	2	ABY37143	Ab37143 Chlamydia
1291	6	5.1	74	5	ABP09615	Abp09615 Human ORF	1364	6	5.1	110	4	ABG17905	Abg17905 Novel hum
1292	6	5.1	76	4	AAW84236	Aaw84236 Human imm	1365	6	5.1	110	9	ADY53933	Ady53933 Corn prot
1293	6	5.1	77	4	AAW90630	Aaw90630 Human imm	1366	6	5.1	113	4	AAO07742	Aao07742 Human pol
1294	6	5.1	77	4	AAU59090	Aau59090 Propionib	1367	6	5.1	113	4	ABG23001	Abg23001 Novel hum
1295	6	5.1	77	6	ABM55609	Abm55609 Propionib	1368	6	5.1	114	5	ABP33635	Abp33635 Human ORF
1296	6	5.1	78	9	ABM93937	Abm93937 M. xanthu	1369	6	5.1	115	4	AAW06666	Aaw06666 Human foe
1297	6	5.1	79	4	AAW31697	Aaw31697 Antigenic	1370	6	5.1	115	6	ABU01952	Abu01952 S. pneumo
1298	6	5.1	79	4	ABG04784	Abg04784 Novel hum	1371	6	5.1	115	8	ADK46883	Adk46883 Streptoco
1299	6	5.1	80	2	AAR12755	Aar12755 Part of N	1372	6	5.1	117	4	ABG14245	Abg14245 Novel hum
1300	6	5.1	80	2	AAR12753	Aar12753 Part of N	1373	6	5.1	117	5	ABW77828	Abw77828 Amino aci
1301	6	5.1	80	2	AAR12752	Aar12752 Part of N	1374	6	5.1	118	4	AGW73501	Agw73501 Human sec
1302	6	5.1	80	2	AAR12756	Aar12756 Part of N	1375	6	5.1	118	4	ABG02037	Abg02037 Novel hum
1303	6	5.1	80	2	AAR12754	Aar12754 Part of N	1376	6	5.1	119	3	ABP39397	Abp39397 Human sec
1304	6	5.1	80	2	AAR24404	Aar24404 Human C h	1377	6	5.1	120	4	AAU16289	Aau16289 Human nov
1305	6	5.1	80	2	AAR24403	Aar24403 Human C h	1378	6	5.1	120	4	AAO02910	Aao02910 Human pol
1306	6	5.1	80	2	AAR24406	Aar24406 Human C h	1379	6	5.1	120	4	ABG08126	Abg08126 Novel hum
1307	6	5.1	80	2	AAR24402	Aar24402 Human C h	1380	6	5.1	120	6	ABU55358	Abu55358 Human nov
1308	6	5.1	80	2	AAR24405	Aar24405 Human C h	1381	6	5.1	120	8	ADR34805	Adr34805 Novel S.
1309	6	5.1	80	7	ADA49413	Ada49413 Multi-epi	1382	6	5.1	120	9	AEA58675	Aea58675 Streptoco
1310	6	5.1	80	8	ADO24091	Ado24091 Epigene c	1383	6	5.1	122	2	AAW585965	Aaw58596 S. pneumo
1311	6	5.1	80	9	ADZ40593	Adz40593 HCV.3a2 (-	1384	6	5.1	122	6	ADB11104	Adb11104 Alloiococ
1312	6	5.1	81	6	ABM15911	Abm15911 Mycobacte	1385	6	5.1	124	3	AAG33339	Ag33339 Zea may
1313	6	5.1	83	4	ABM21622	Abm21622 Peptide #	1386	6	5.1	124	8	ADW1723	Adw1723 Plant ful
1314	6	5.1	83	4	ABB43982	Abb43982 Peptide #	1387	6	5.1	125	4	ABW12174	Abw12174 Human sec
1315	6	5.1	83	4	AAW37924	Aaw37924 Peptide #	1388	6	5.1	125	5	ABP00650	Abp00650 Human ORF
1316	6	5.1	83	4	ABB26891	Abb26891 Protein #	1389	6	5.1	125	8	ADW78896	Adw78896 Plant ful
1317	6	5.1	83	4	AAW77707	Aaw77707 Human bon	1390	6	5.1	126	3	AAG33338	Ag33338 Zea may
1318	6	5.1	83	4	AAW64985	Aaw64985 Human bra	1391	6	5.1	127	2	AAW35301	Aaw35301 Human pol
1319	6	5.1	83	4	ABG59362	Abg59362 Human liv	1392	6	5.1	127	2	ABW50267	Abw50267 Polymaras
1320	6	5.1	83	5	ABG46738	Abg46738 Human pep	1393	6	5.1	127	6	ABR47556	Abra47556 Breast ca
1321	6	5.1	83	6	ABR56729	AbR56729 Human sec	1394	6	5.1	127	7	ADF76608	Adf76608 Novel hum
1322	6	5.1	86	2	AAW09051	Aaw09051 Hepatitis	1395	6	5.1	127	7	ADM85791	Adm85791 Human pro
1323	6	5.1	86	4	ABG20560	Abg20560 Novel hum	1396	6	5.1	127	8	ADL91494	Adl91494 Human imm
1324	6	5.1	87	2	AAW49652	Aaw49652 HCV pepti	1397	6	5.1	127	8	ADN04676	Adn04676 Anti-peori
1325	6	5.1	95	6	ABU28008	Abu28008 Protein e	1398	6	5.1	127	8	ADN98783	Adn98783 Novel hum
1326	6	5.1	96	4	AAO09002	Aao09002 Human pol	1399	6	5.1	127	8	ADP55579	Adp55579 Human PRO
1327	6	5.1	96	5	ABP29140	Abp29140 Streptoco	1400	6	5.1	127	8	ADP55112	Adp55112 Human PRO
1328	6	5.1	96	9	ABB42456	Abb42456 L. pneumo	1401	6	5.1	127	8	ADU06509	Adu06509 Novel bro
1329	6	5.1	97	7	ADC94028	Adc94028 E. faeciu	1402	6	5.1	127	9	ABM95146	Abm95146 M. xanthu
1330	6	5.1	97	7	ADN11795	Adn11795 c-Src SH2	1403	6	5.1	128	3	AAW40273	Aaw40273 Human ORF
1331	6	5.1	98	2	AAW59763	Aaw59763 Amino aci	1404	6	5.1	128	5	ABP01463	Abp01463 Human ORF
1332	6	5.1	98	2	AAW59756	Aaw59756 Amino aci	1405	6	5.1	128	5	ABP64800	Abp64800 Human pro
1333	6	5.1	98	2	AAW59758	Aaw59758 Amino aci	1406	6	5.1	129	2	AAW04871	Aaw04871 Mycobacte
1334	6	5.1	98	2	AAW59759	Aaw59759 Amino aci	1407	6	5.1	129	7	ADH86015	Adh86015 Enterococ
1335	6	5.1	98	2	AAW59760	Aaw59760 Amino aci	1408	6	5.1	130	2	AAW03981	Aaw03981 DET1-DET2
1336	6	5.1	98	2	AAW59757	Aaw59757 Amino aci	1409	6	5.1	130	2	AAW02119	Aaw02119 DET1-DET2
1337	6	5.1	98	4	AAU31683	Aau31683 Novel hum	1410	6	5.1	130	2	AAW11285	Aaw11285 DET1-DET2
1338	6	5.1	99	5	ABP53491	Abp53491 V-Src SH2	1411	6	5.1	130	2	AAW19623	Aaw19623 Human src

1412	6	5.1	130	4	AAM83255	Human imm	1485	6	5.1	153	4	AAB88423	Aab88423 Human mem
1413	6	5.1	130	4	ABG23720	Novel hum	1486	6	5.1	153	7	ADC00265	Adc00265 Enterohae
1414	6	5.1	130	4	ADA55371	Human pro	1487	6	5.1	153	8	ADN99251	Adn99251 Novel hum
1415	6	5.1	130	8	ADX68201	Plant ful	1488	6	5.1	153	8	ADX68668	Adx68668 Plant ful
1416	6	5.1	131	2	AAV73815	Human pro	1489	6	5.1	153	9	ADY63209	Ady63209 Human clo
1417	6	5.1	132	5	ABF63705	Human pan	1490	6	5.1	154	3	AGS33337	AgS33337 Zea mayS
1418	6	5.1	132	5	ABP63705	Human ORP	1491	6	5.1	154	8	ADK01698	AdK01698 Hepatitis
1419	6	5.1	132	8	ADX80418	Plant ful	1492	6	5.1	155	4	AAB68931	Aab68931 Neisseria
1420	6	5.1	133	3	AAV94956	Human sec	1493	6	5.1	157	4	AAU22988	Aau22988 Novel hum
1421	6	5.1	133	4	AAU60129	Propionib	1494	6	5.1	157	7	ADA49433	Ada49433 Multi-epi
1422	6	5.1	133	6	ADA54198	Human pro	1495	6	5.1	157	7	ADP58966	Adp58966 Human pol
1423	6	5.1	133	6	ARM56648	Propionib	1496	6	5.1	157	7	ABO75397	AbO75397 Pseudomon
1424	6	5.1	133	8	ADY04853	Plant ful	1497	6	5.1	157	8	ADO24111	AdO24111 Epigene c
1425	6	5.1	134	4	AAW74189	Human bon	1498	6	5.1	157	8	ADX76746	AdX76746 Plant ful
1426	6	5.1	134	4	ABG55982	Human liv	1499	6	5.1	157	9	ADZ40613	AdZ40613 PfCTL.2 m
1427	6	5.1	134	5	ABG44134	Human pep	1500	6	5.1	158	7	ABM88481	Abm88481 Rice abio
1428	6	5.1	134	8	ADY11719	Plant ful	1501	6	5.1	158	7	ADM04428	Adm04428 Human pro
1429	6	5.1	134	8	ADX67938	Plant ful	1502	6	5.1	159	7	ADM04428	Adm04428 Human pro
1430	6	5.1	135	8	ADY04808	Plant ful	1503	6	5.1	160	6	ABJ38629	AbJ38629 Human TIG
1431	6	5.1	135	8	ADY09953	Plant ful	1504	6	5.1	160	9	ADW43852	AdW43852 Rat prepr
1432	6	5.1	135	8	ADX67942	Plant ful	1505	6	5.1	161	2	AAV59955	Aav59955 Human end
1433	6	5.1	136	2	AAV42641	S. typhim	1506	6	5.1	161	3	AG47878	Ag47878 Arabidops
1434	6	5.1	136	6	ABU34161	Protein e	1507	6	5.1	161	3	AG49272	Ag49272 Arabidops
1435	6	5.1	136	8	ADX67115	Plant ful	1508	6	5.1	161	3	AG16350	Ag16350 Arabidops
1436	6	5.1	137	5	ABP43529	Human sec	1509	6	5.1	161	3	AG24071	Ag24071 Arabidops
1437	6	5.1	137	8	ADX88700	Plant ful	1510	6	5.1	161	3	AG49269	Ag49269 Arabidops
1438	6	5.1	138	8	ADX95661	Plant ful	1511	6	5.1	161	7	ABO75056	AbO75056 Pseudomon
1439	6	5.1	138	9	ABM95557	M. xanthu	1512	6	5.1	161	8	ADX68749	AdX68749 Plant ful
1440	6	5.1	139	3	AAV75491	Neisseria	1513	6	5.1	162	1	RAP51201	Rap51201 Antigenic
1441	6	5.1	139	3	AAV75514	Neisseria	1514	6	5.1	162	7	ADC27577	AdC27577 Human col
1442	6	5.1	139	3	AG40877	Zea mayS	1515	6	5.1	163	3	AG49268	Ag49268 Arabidops
1443	6	5.1	139	6	ABU34465	Protein e	1516	6	5.1	163	3	AG16349	Ag16349 Arabidops
1444	6	5.1	139	6	ABU36812	Plant ful	1517	6	5.1	163	3	AG49271	Ag49271 Arabidops
1445	6	5.1	139	8	ADX67931	Plant ful	1518	6	5.1	163	3	AG49271	Ag49271 Arabidops
1446	6	5.1	139	8	ADX92979	Plant ful	1519	6	5.1	163	3	AG32274	Ag32274 Arabidops
1447	6	5.1	140	4	AAO01783	Human pol	1520	6	5.1	163	3	AG47877	Ag47877 Arabidops
1448	6	5.1	140	7	ADF18732	DNA-depen	1521	6	5.1	163	3	AG24070	Ag24070 Arabidops
1449	6	5.1	141	4	AAU16288	Human nov	1522	6	5.1	163	8	ABO60002	AbO60002 Human gen
1450	6	5.1	141	6	ADA54699	Human pro	1523	6	5.1	163	9	ADZ69261	AdZ69261 Rat adipo
1451	6	5.1	141	6	ADA55170	Human pro	1524	6	5.1	164	2	AAV37563	Aav37563 Amino aci
1452	6	5.1	141	6	ABU55357	Human nov	1525	6	5.1	164	2	ADP66077	Adp66077 Hepatitis
1453	6	5.1	141	8	ADO22008	Rhodobact	1526	6	5.1	164	4	ABG21786	AbG21786 Novel hum
1454	6	5.1	142	4	AAO05704	Human pol	1527	6	5.1	164	8	ADI34119	AdI34119 Ketogulon
1455	6	5.1	142	5	ADK34520	Novel hum	1528	6	5.1	165	5	ABB72314	Abb72314 Rat prote
1456	6	5.1	142	6	ABU23423	Protein e	1529	6	5.1	165	8	ADT59536	Adt59536 Plant pol
1457	6	5.1	142	8	ADT87268	Plant ful	1530	6	5.1	165	7	ABO80939	AbO80939 Pseudomon
1458	6	5.1	143	3	AB41344	Human ORP	1531	6	5.1	166	2	AAV36400	Aav36400 Fragment
1459	6	5.1	143	4	AB81187	Human rev	1532	6	5.1	168	6	AAO26582	Aao26582 Alternari
1460	6	5.1	143	5	ABP35424	Human tra	1533	6	5.1	168	6	AAO26581	Aao26581 Alternari
1461	6	5.1	143	6	ADA33628	Acinetoba	1534	6	5.1	169	3	AG25717	Ag25717 Arabidops
1462	6	5.1	143	8	ADX71772	Plant ful	1535	6	5.1	169	8	ADN46924	Adn46924 Thermococ
1463	6	5.1	144	4	AAW83990	Human imm	1536	6	5.1	170	6	ABP57467	Abp57467 Mycobacte
1464	6	5.1	144	6	ADA35983	Acinetoba	1537	6	5.1	170	6	ABU34463	Abu34463 Protein e
1465	6	5.1	144	8	ADT58104	Plant pol	1538	6	5.1	170	6	ABU36814	Abu36814 Protein e
1466	6	5.1	145	7	ABM89247	Rice abio	1539	6	5.1	170	6	ABM73538	Abm73538 Staphyloc
1467	6	5.1	145	8	ADX90247	Plant ful	1540	6	5.1	170	8	ADK48153	AdK48153 Streptoco
1468	6	5.1	145	8	ADX93828	Plant ful	1541	6	5.1	172	2	AAV00230	Aav00230 Enterococ
1469	6	5.1	145	8	ADX75006	Plant ful	1542	6	5.1	172	5	ABP43449	Abp43449 E faecali
1470	6	5.1	146	3	AG41182	Zea mayS	1543	6	5.1	172	6	ABU88477	Abu88477 E. faecal
1471	6	5.1	146	6	ABP78075	N. gonorr	1544	6	5.1	172	6	ABU29433	Abu29433 Protein e
1472	6	5.1	146	8	ADY25286	Plant ful	1545	6	5.1	172	6	ABU13728	Abu13728 Enterococ
1473	6	5.1	146	9	ABM92872	M. xanthu	1546	6	5.1	172	7	ABO78044	AbO78044 Pseudomon
1474	6	5.1	148	3	AAV73332	HTM clon	1547	6	5.1	172	9	ADY39176	Ady39176 Novel Ent
1475	6	5.1	150	4	ABG26860	Novel hum	1548	6	5.1	173	2	AAW89770	Aaw89770 Staphyloc
1476	6	5.1	150	4	ABG27269	Novel hum	1549	6	5.1	173	4	AAU23455	Aau23455 Novel hum
1477	6	5.1	150	8	ADS14928	Pseudomon	1550	6	5.1	174	4	AAW96395	Aaw96395 Putative
1478	6	5.1	150	8	ADY10619	Plant ful	1551	6	5.1	174	9	ADZ22188	AdZ22188 Full leng
1479	6	5.1	150	8	ADY10657	Plant ful	1552	6	5.1	175	3	AG25716	Ag25716 Arabidops
1480	6	5.1	151	7	ABO83313	Pseudomon	1553	6	5.1	176	8	ADX88225	Adx88225 Plant ful
1481	6	5.1	152	3	ABG41181	Zea mayS	1554	6	5.1	176	9	ABM94808	Abm94808 M. xanthu
1482	6	5.1	152	6	ABP79124	N. gonorr	1555	6	5.1	179	5	AAW49456	Aaw49456 Phormidiu
1483	6	5.1	152	6	ABU35673	Protein e	1556	6	5.1	179	8	ADY04453	Ady04453 Plant ful
1484	6	5.1	152	8	ADR20454	Human EVE	1557	6	5.1	180	4	AAU20510	Aau20510 Human sec

1558	6	5.1	180	4	ABO3072	Human exp	1631	6	5.1	214	4	ABG23521	Novel hum
1559	6	5.1	180	8	ADK68733	Plant ful	1632	6	5.1	214	6	ABU21350	Protein e
1560	6	5.1	181	5	ABBS4325	Lactococc	1633	6	5.1	214	7	ADM04666	Human pro
1561	6	5.1	181	8	ADS24804	Bacterial	1634	6	5.1	216	8	ABO84901	Human can
1562	6	5.1	182	6	AAO26580	Alternari	1635	6	5.1	217	4	ABG08127	Novel hum
1563	6	5.1	182	8	ADY09244	Plant ful	1636	6	5.1	217	8	ADG95104	Human the
1564	6	5.1	182	9	AEA23177	Pseudomon	1637	6	5.1	217	8	ADT58263	Plant pol
1565	6	5.1	183	6	ABU22704	Protein e	1638	6	5.1	217	8	ADV67961	Biologica
1566	6	5.1	183	8	ADY10644	Plant ful	1639	6	5.1	218	3	AGS4892	Arabidops
1567	6	5.1	184	7	ABM89930	Rice abio	1640	6	5.1	218	7	ABO71185	Pseudomon
1568	6	5.1	185	8	ADS24907	Bacterial	1641	6	5.1	218	8	ADS24724	Bacterial
1569	6	5.1	186	6	ABU01103	S. pneumo	1642	6	5.1	220	4	ABG07806	Novel hum
1570	6	5.1	186	7	ADH86292	Enterococ	1643	6	5.1	221	3	AAV57897	Human tra
1571	6	5.1	186	7	ABO81333	Pseudomon	1644	6	5.1	221	3	AAV87080	Human sec
1572	6	5.1	186	8	ADK46312	Streptoco	1645	6	5.1	221	4	ABH8606	Human hyd
1573	6	5.1	186	8	ADM92123	S. pneumon	1646	6	5.1	221	4	AAW79069	Human pro
1574	6	5.1	186	8	ADS24765	Bacterial	1647	6	5.1	221	4	AAE06057	Human gen
1575	6	5.1	186	8	ADT60342	Plant pol	1648	6	5.1	221	4	ABH88441	Human mem
1576	6	5.1	187	3	AAV81611	Streptoco	1649	6	5.1	221	4	AGS92353	C glutami
1577	6	5.1	188	4	ABBS6901	Drosophil	1650	6	5.1	221	4	ABG07522	Novel hum
1578	6	5.1	188	7	ABO66828	Klebsiell	1651	6	5.1	221	5	ADA43378	Human ast
1579	6	5.1	190	5	AAE18631	Equine hi	1652	6	5.1	221	5	ABB08267	Human pol
1580	6	5.1	190	6	ADA33262	Acinetoba	1653	6	5.1	221	5	ABP61429	Human NF-
1581	6	5.1	191	4	ABG13980	Novel hum	1654	6	5.1	221	5	ABG33879	Human sec
1582	6	5.1	191	8	ADK72255	Plant ful	1655	6	5.1	221	5	ABU02855	S. pneumo
1583	6	5.1	192	4	ABG14247	Novel hum	1656	6	5.1	221	6	ADA57061	Human sec
1584	6	5.1	192	8	ADY04833	Plant ful	1657	6	5.1	221	6	ADA40914	Human sec
1585	6	5.1	193	7	ADH86544	Enterococ	1658	6	5.1	221	6	ABR47827	Human sec
1586	6	5.1	193	8	ADY72826	Plant ful	1659	6	5.1	221	6	ABR41028	Human MAP
1587	6	5.1	193	8	ADY22892	Plant ful	1660	6	5.1	221	7	ABR84649	Human SGA
1588	6	5.1	194	6	ABM70654	Phototrab	1661	6	5.1	221	8	ADK48501	Streptoco
1589	6	5.1	194	8	ADK71176	Plant ful	1662	6	5.1	221	8	ADN92256	Novel hum
1590	6	5.1	195	4	AAU20588	Human sec	1663	6	5.1	221	9	ADY63245	Human clo
1591	6	5.1	195	4	ABG10071	Novel hum	1664	6	5.1	222	3	AAV81586	Streptoco
1592	6	5.1	196	4	AAV90945	C glutami	1665	6	5.1	222	8	ABO60005	Human gen
1593	6	5.1	197	5	AAO16536	Histone a	1666	6	5.1	223	6	ABU19948	Protein e
1594	6	5.1	198	2	AAV39318	ORF1 pro	1667	6	5.1	223	8	ADT57688	Plant pol
1595	6	5.1	200	5	ABB90263	Human pol	1668	6	5.1	224	4	ABE64303	Drosophil
1596	6	5.1	200	7	ADF18729	DNA-depen	1669	6	5.1	224	8	ADN89486	A. oryzae
1597	6	5.1	200	8	ADR86353	Aspergill	1670	6	5.1	225	2	AAW01112	Xylanase
1598	6	5.1	201	2	AAV27061	Recombina	1671	6	5.1	225	2	AAW05187	Endo-1,4-
1599	6	5.1	202	8	ADL04728	M. catarr	1672	6	5.1	225	7	ABR63119	Thermomyc
1600	6	5.1	202	8	ADR94509	Novel S.	1673	6	5.1	225	8	ADK79059	Plant ful
1601	6	5.1	202	8	ADY09246	Plant ful	1674	6	5.1	225	9	ABE00305	Xylanase
1602	6	5.1	202	9	AEA58379	Streptoco	1675	6	5.1	226	3	AAG42960	Arabidops
1603	6	5.1	203	3	AAO6232	Arabidops	1676	6	5.1	226	7	ADJ69171	Human hea
1604	6	5.1	203	7	ABO75289	Pseudomon	1677	6	5.1	226	7	ABO68982	Pseudomon
1605	6	5.1	203	8	ABO59518	Human gen	1678	6	5.1	226	8	ADY09505	Plant ful
1606	6	5.1	204	2	AAU20610	C10-11 NA	1679	6	5.1	227	2	AAV00247	Enterococ
1607	6	5.1	205	4	AAU20386	Human sec	1680	6	5.1	227	5	ABP43466	E faecali
1608	6	5.1	205	4	ABBI17052	Human ner	1681	6	5.1	227	6	ABU88494	E. faecoc
1609	6	5.1	206	6	ABU23285	Protein e	1682	6	5.1	227	6	ABU13745	Enterococ
1610	6	5.1	206	6	ABU22243	Protein e	1683	6	5.1	227	8	ADY08598	Plant ful
1611	6	5.1	207	6	AAU33541	Klebsiell	1684	6	5.1	227	9	ADY39210	Novel Ent
1612	6	5.1	207	6	ABU37667	Protein e	1685	6	5.1	228	4	ABE64950	Gene 11 h
1613	6	5.1	207	9	ADK39793	HIV Nef p	1686	6	5.1	229	7	ABO81883	Pseudomon
1614	6	5.1	207	9	ADK39865	HIV Nef p	1687	6	5.1	230	7	ADC97588	E. faeciu
1615	6	5.1	208	4	AAU36059	Klebsiell	1688	6	5.1	231	6	ADA36641	Acinetoba
1616	6	5.1	209	2	AAU20612	C10-14 NA	1689	6	5.1	232	3	AAG42301	Arabidops
1617	6	5.1	209	2	AAV59784	Human nor	1690	6	5.1	232	4	ABG12545	Novel hum
1618	6	5.1	209	7	ADM26718	Hyperther	1691	6	5.1	232	8	ADN99515	Novel hum
1619	6	5.1	210	4	AAO06126	Human pol	1692	6	5.1	232	8	ADK66541	Plant ful
1620	6	5.1	211	3	AAV51675	Murine cl	1693	6	5.1	233	9	ADR96478	Novel S.
1621	6	5.1	211	5	ADL16801	Murine NO	1694	6	5.1	233	9	AEA60348	Streptoco
1622	6	5.1	211	5	ADL16791	Murine NO	1695	6	5.1	235	3	AAO62231	Arabidops
1623	6	5.1	211	6	ABM15875	Mycobacte	1696	6	5.1	235	4	AAW25470	Human pro
1624	6	5.1	211	8	ADJ76259	Marker ge	1697	6	5.1	236	2	AAV27060	Equine Fc
1625	6	5.1	211	8	ADJ76221	Marker ge	1698	6	5.1	237	2	AAV04875	Mycobacte
1626	6	5.1	211	8	ADN99663	Novel hum	1699	6	5.1	237	8	ADU16800	M. tuberc
1627	6	5.1	211	8	ADK93739	Plant ful	1700	6	5.1	238	6	ABU22091	Protein e
1628	6	5.1	212	4	ABG13284	Novel hum	1701	6	5.1	238	7	ABO74641	Pseudomon
1629	6	5.1	213	4	AAV92445	C glutami	1702	6	5.1	238	9	ABM93188	M. xanthu
1630	6	5.1	214	2	AAV29943	Zea mays	1703	6	5.1	239	3	AAV42300	Arabidops

1704	6	5.1	239	5	ABP41821	Abp41821 Human ova	1777	6	5.1	264	6	ADB10584	Adb10584 Alloiooc
1705	6	5.1	239	5	AB54226	LaCococcc	1778	6	5.1	264	6	ADN25533	Bacterial
1706	6	5.1	239	6	ADB09538	Alloiooccc	1779	6	5.1	264	9	ABE13169	C. glutam
1707	6	5.1	239	8	ADJ79496	Epstein-B	1780	6	5.1	265	8	ABU35807	Protein e
1708	6	5.1	239	8	ADJ88356	Plant ful	1781	6	5.1	265	8	ADN26509	Bacterial
1709	6	5.1	240	7	ABO79366	Pseudomon	1782	6	5.1	265	8	ADN26750	Bacterial
1710	6	5.1	240	8	ADS28796	Bacterial	1783	6	5.1	265	8	ADT60937	Plant pol
1711	6	5.1	240	8	ADY07284	Plant ful	1784	6	5.1	266	4	AAAB79943	Corynebac
1712	6	5.1	241	6	ABU37809	Protein e	1785	6	5.1	266	4	AAAG92701	C glutami
1713	6	5.1	242	3	ABAB43185	Human ORF	1786	6	5.1	266	8	ADS42405	Bacterial
1714	6	5.1	242	4	ABAB94286	Human pro	1787	6	5.1	266	8	ADY05685	Plant ful
1715	6	5.1	242	5	ABAB89567	Human pol	1788	6	5.1	266	8	ADX67795	Plant ful
1716	6	5.1	242	8	ADS28902	Bacterial	1789	6	5.1	267	5	AAU91162	Bordellia
1717	6	5.1	243	7	ABO70893	Pseudomon	1790	6	5.1	267	7	ABD74365	Mycobacte
1718	6	5.1	243	8	ADY06386	Plant ful	1791	6	5.1	267	7	ABO83578	Pseudomon
1719	6	5.1	244	5	ABP27420	Streptoco	1792	6	5.1	267	8	ABO60468	Human gen
1720	6	5.1	245	2	AAV04898	Mycobacte	1793	6	5.1	267	4	AAU49111	Propionib
1721	6	5.1	245	8	ADJ79592	Hepatitis	1794	6	5.1	268	6	ABM45630	Propionib
1722	6	5.1	245	8	ADU16885	M. tuberc	1795	6	5.1	268	4	ADG22351	Cyanophag
1723	6	5.1	245	8	ADY10757	Plant ful	1796	6	5.1	269	4	AAU66854	Propionib
1724	6	5.1	246	3	AAAB52085	Gene 33 h	1797	6	5.1	269	6	ABM63373	Propionib
1725	6	5.1	246	6	ABM64672	Propionib	1798	6	5.1	269	7	AAO24040	Aspergill
1726	6	5.1	246	7	ABO78865	Pseudomon	1799	6	5.1	269	8	ADN21926	Bacterial
1727	6	5.1	246	8	ADY22394	Plant ful	1800	6	5.1	269	8	ADN24683	Bacterial
1728	6	5.1	246	8	ADY09476	Plant ful	1801	6	5.1	269	8	ADU63802	A. oryzae
1729	6	5.1	247	4	AAU62596	Propionib	1802	6	5.1	269	9	ADY09187	Plant ful
1730	6	5.1	247	4	ABG13024	Novel hum	1803	6	5.1	269	9	ADW23380	Aspergill
1731	6	5.1	247	6	ABM59115	Propionib	1804	6	5.1	270	7	ABD80163	Mycobacte
1732	6	5.1	247	7	ADM05413	Human pro	1805	6	5.1	271	4	ABM66816	Drosophil
1733	6	5.1	248	2	AAV04876	Mycobacte	1806	6	5.1	271	4	ABG03898	Novel hum
1734	6	5.1	248	3	AGI14631	Arabidops	1807	6	5.1	271	7	ABO69792	Pseudomon
1735	6	5.1	248	8	ADI37324	M. tuberc	1808	6	5.1	271	8	ADS22212	Bacterial
1736	6	5.1	248	8	ADU16802	M. tuberc	1809	6	5.1	273	4	ABG17593	Novel hum
1737	6	5.1	249	3	AGO7675	Arabidops	1810	6	5.1	273	4	ABG14246	Novel hum
1738	6	5.1	250	7	ABE08522	Novel pro	1811	6	5.1	273	7	ADF18731	DNA-depen
1739	6	5.1	251	3	AAV44450	Mutant ch	1812	6	5.1	274	6	ABJ26874	Immune-re
1740	6	5.1	251	4	ABG13152	Novel hum	1813	6	5.1	274	6	ABM15858	Mycobacte
1741	6	5.1	251	8	ADX78164	Plant ful	1814	6	5.1	274	6	ABU49384	Protein e
1742	6	5.1	252	4	ABG09030	Novel hum	1815	6	5.1	275	4	AAAB73667	Mouse age
1743	6	5.1	252	5	ABB49400	Listeria	1816	6	5.1	275	6	ABU01362	S. pneumo
1744	6	5.1	253	4	AAU59879	Propionib	1817	6	5.1	275	7	ABO61935	Klebsiell
1745	6	5.1	253	6	ABM56398	Propionib	1818	6	5.1	275	8	ADM90948	Human pha
1746	6	5.1	253	7	ABO75616	Pseudomon	1819	6	5.1	276	7	ADA49457	Multi-epi
1747	6	5.1	254	4	AAAG92456	C glutami	1820	6	5.1	276	8	ADQ24135	Epigene c
1748	6	5.1	254	6	ABU06072	N. mening	1821	6	5.1	276	9	ADZ40637	pFHTL mul
1749	6	5.1	254	7	ADG73395	E. faecali	1822	6	5.1	277	4	ABG04213	Novel hum
1750	6	5.1	254	7	ADG73397	E. faecali	1823	6	5.1	277	6	ABU17212	Protein e
1751	6	5.1	254	9	AEC10571	Enterococ	1824	6	5.1	277	7	ADC39208	Novel hum
1752	6	5.1	254	9	AEC10573	Enterococ	1825	6	5.1	277	4	ABG12548	Novel hum
1753	6	5.1	255	2	AAV27058	Equine Fc	1826	6	5.1	279	4	AAAB94900	Human pro
1754	6	5.1	255	9	ABM94181	M. xanthu	1827	6	5.1	280	5	ABP29632	Streptoco
1755	6	5.1	256	6	ABM70498	Phototrab	1828	6	5.1	280	6	ABU01399	S. pneumo
1756	6	5.1	257	5	AAU74649	Oestrogen	1829	6	5.1	280	6	ABP81380	Streptoco
1757	6	5.1	257	8	ABU33688	Protein e	1830	6	5.1	280	7	ADC31428	Human nov
1758	6	5.1	258	7	ABO71514	Pseudomon	1831	6	5.1	280	8	ADK48692	Streptoco
1759	6	5.1	259	2	AAW20187	H. pylori	1832	6	5.1	280	8	ADM92149	S. pneumon
1760	6	5.1	259	5	ABP28533	Streptoco	1833	6	5.1	280	8	ADN46709	Thermococ
1761	6	5.1	259	5	ABP28532	Streptoco	1834	6	5.1	280	8	ADX77818	Plant ful
1762	6	5.1	259	8	ADV88598	Streptoco	1835	6	5.1	281	3	AAAG42959	Arabidops
1763	6	5.1	259	8	ADV79851	Streptoco	1836	6	5.1	281	7	ADA49407	Multi-epi
1764	6	5.1	259	8	ADV82004	Streptoco	1837	6	5.1	281	8	ADO24085	Epigene c
1765	6	5.1	261	6	ABU36639	Protein e	1838	6	5.1	281	9	ADZ40587	HCV-2 mul
1766	6	5.1	261	6	ABU34460	Protein e	1839	6	5.1	282	4	ABG25874	Novel hum
1767	6	5.1	261	7	ABO71863	Pseudomon	1840	6	5.1	282	7	ABD70058	C. neofo
1768	6	5.1	262	5	ABG80026	A. cryptu	1841	6	5.1	282	7	ABO65724	Klebsiell
1769	6	5.1	262	5	ABB80942	S. coelic	1842	6	5.1	283	4	ABB62426	Drosophil
1770	6	5.1	262	6	ABO809540	Alloiooc	1843	6	5.1	283	6	ABU15173	Protein e
1771	6	5.1	262	7	ADC01009	Enterohae	1844	6	5.1	283	7	ABO62703	Klebsiell
1772	6	5.1	262	7	ABE14793	Streptomy	1845	6	5.1	284	4	ABG04337	Novel hum
1773	6	5.1	263	8	ADH39694	Streptomy	1846	6	5.1	284	6	ABU22112	Protein e
1774	6	5.1	263	8	ADS22803	Bacterial	1847	6	5.1	284	6	ABU25168	Protein e
1775	6	5.1	264	4	ABB61745	Drosophil	1848	6	5.1	284	8	ADX73849	Plant ful
1776	6	5.1	264	4	AAAG91869	C glutami	1849	6	5.1	285	2	AAV73976	Human pro

1850 6 5.1 285 4 ABG06018 Novel hum
1851 6 5.1 285 5 ABG93284 C. albica
1852 6 5.1 286 3 AAG49695 Arabidops
1853 6 5.1 286 3 AAG49695 Arabidops
1854 6 5.1 286 7 ADJ68620 Human hea
1855 6 5.1 286 9 ADY70346 Human bet
1856 6 5.1 287 6 ADA35809 Acinetoba
1857 6 5.1 287 8 ADL04791 M. catarr
1858 6 5.1 287 9 ABM94019 M. xanthu
1859 6 5.1 288 3 AAG49694 Arabidops
1860 6 5.1 288 3 AAG14630 Arabidops
1861 6 5.1 288 4 ABG28610 Novel hum
1862 6 5.1 288 4 AAU69506 Human pur
1863 6 5.1 288 7 ABO79233 Pseudomon
1864 6 5.1 288 8 ADS22563 Bacterial
1865 6 5.1 288 8 ADS25850 Bacterial
1866 6 5.1 289 2 AAU00249 Enterococ
1867 6 5.1 289 4 AAU20512 Human sec
1868 6 5.1 289 4 ABG23002 Novel hum
1869 6 5.1 289 5 ABP43468 E faecali
1870 6 5.1 289 6 ABU88496 E. faecal
1871 6 5.1 289 6 ABU13747 Enterococ
1872 6 5.1 289 9 ADY39214 Novel Ent
1873 6 5.1 290 3 AAG25199 Arabidops
1874 6 5.1 290 6 ADA36103 Acinetoba
1875 6 5.1 290 7 ABM87938 Rice abio
1876 6 5.1 291 4 AAM180327 Peptide #
1877 6 5.1 291 4 ABB37361 Peptide #
1878 6 5.1 291 4 ABB32109 Peptide #
1879 6 5.1 291 4 ABB22647 Protein #
1880 6 5.1 291 4 AAM70491 Human bon
1881 6 5.1 291 4 ABG52172 Human liv
1882 6 5.1 291 4 AAM05934 Peptide #
1883 6 5.1 291 5 ABG40130 Human pep
1884 6 5.1 291 7 ABO74080 Pseudomon
1885 6 5.1 292 3 AAG06230 Arabidops
1886 6 5.1 292 4 AAB80103 Corynebac
1887 6 5.1 292 7 ADM25467 Hyperther
1888 6 5.1 293 4 ABG04338 Novel hum
1889 6 5.1 293 3 AAG42299 Arabidops
1890 6 5.1 294 4 ABG17979 Novel hum
1891 6 5.1 294 8 ADS25255 Bacterial
1892 6 5.1 294 8 ADS25533 Bacterial
1893 6 5.1 295 2 AAW58541 Aspergill
1894 6 5.1 295 3 AAG25198 Arabidops
1895 6 5.1 295 3 AAG49693 Arabidops
1896 6 5.1 295 4 AAB73666 Mouse age
1897 6 5.1 295 4 ABG19113 Novel hum
1898 6 5.1 295 4 ABG14036 Novel hum
1899 6 5.1 295 8 ADR43134 IPT-like
1900 6 5.1 296 3 AAG08463 Arabidops
1901 6 5.1 296 3 AAG52192 Arabidops
1902 6 5.1 296 5 ABP61056 Lactobaci
1903 6 5.1 296 6 ABU39672 Protein e
1904 6 5.1 296 7 ADE12772 L. rhamno
1905 6 5.1 297 3 AAY96574 HSV-1 VP2
1906 6 5.1 297 5 ABP66101 Bifidobac
1907 6 5.1 297 6 ABU23021 Protein e
1908 6 5.1 297 8 ADN21998 Bacterial
1909 6 5.1 298 8 ADN24756 Bacterial
1910 6 5.1 298 2 AAW35300 Human isl
1911 6 5.1 299 7 ABO74089 Pseudomon
1912 6 5.1 299 8 ADM90917 Human pha
1913 6 5.1 299 8 ADM90915 Human pha
1914 6 5.1 300 3 AAG08462 Arabidops
1915 6 5.1 300 3 AAG52191 Arabidops
1916 6 5.1 300 4 AAU33689 Pseudomon
1917 6 5.1 300 6 ABU15693 Protein e
1918 6 5.1 300 6 ABU39771 Protein e
1919 6 5.1 300 8 AAG14629 Arabidops
1920 6 5.1 302 3 AAG14629 Arabidops
1921 6 5.1 302 4 ABG10118 Novel hum
1922 6 5.1 302 7 ABO77618 Pseudomon

1923 6 5.1 302 8 ADN99253 Novel hum
1924 6 5.1 302 8 ADR26084 Bacterial
1925 6 5.1 302 8 ADR26431 Bacterial
1926 6 5.1 302 9 AEA02921 Clq-relat
1927 6 5.1 302 9 AEA02929 Clq-relat
1928 6 5.1 303 6 ABU34689 Protein e
1929 6 5.1 303 6 ABU36704 Protein e
1930 6 5.1 303 7 ABM87764 Rice abio
1931 6 5.1 303 8 ADS43339 Bacterial
1932 6 5.1 304 3 AAG07674 Arabidops
1933 6 5.1 304 6 ADA34816 Acinetoba
1934 6 5.1 304 6 ADA34433 Acinetoba
1935 6 5.1 304 7 ABO75280 Pseudomon
1936 6 5.1 304 9 AEB49954 R. etli p
1937 6 5.1 305 4 ABG22875 Novel hum
1938 6 5.1 305 8 ADR28724 Bacterial
1939 6 5.1 305 8 ADS25310 Bacterial
1940 6 5.1 306 2 AAY05766 Streptoco
1941 6 5.1 306 3 AAY91280 Group B S
1942 6 5.1 306 5 ABP28905 Streptoco
1943 6 5.1 306 5 ABP26210 Streptoco
1944 6 5.1 306 7 ABO78052 Pseudomon
1945 6 5.1 306 8 ADS29787 Bacterial
1946 6 5.1 306 8 ADS41717 Bacterial
1947 6 5.1 306 8 ADU69610 S agalact
1948 6 5.1 306 8 ADV88782 Streptoco
1949 6 5.1 306 8 ADV80035 Streptoco
1950 6 5.1 306 8 ADV82173 Streptoco
1951 6 5.1 307 3 AAY56333 Borrelia
1952 6 5.1 307 3 AAY56333 Borrelia
1953 6 5.1 308 4 ABB12341 Human bon
1954 6 5.1 308 5 ABB95179 Human 158
1955 6 5.1 308 5 ABB95187 Human 158
1956 6 5.1 308 7 ADA49421 Multi-epi
1957 6 5.1 308 7 ADA49423 Multi-epi
1958 6 5.1 308 8 ADO24101 Epigene c
1959 6 5.1 308 8 ADO24099 Epigene c
1960 6 5.1 308 9 ADX40834 HCV fusio
1961 6 5.1 308 9 ADZ40603 HCV. 4312 (

1962 6 5.1 308 9 ADZ40601 HCV. 4312 (

1963 6 5.1 308 9 AEA02923 Clq-relat
1964 6 5.1 309 8 ADS21635 Bacterial
1965 6 5.1 310 6 ABU48446 Protein e
1966 6 5.1 311 6 ABP57355 Human sec
1967 6 5.1 312 5 AAO17807 H influen
1968 6 5.1 312 8 ADR43161 IPT-like
1969 6 5.1 312 8 ADR43233 IPT-like
1970 6 5.1 313 4 AAB93705 Human pro
1971 6 5.1 313 6 ABR92073 Human cer
1972 6 5.1 313 8 ADP04897 Sea squir
1973 6 5.1 313 8 ADO20332 Human PRO
1974 6 5.1 313 8 ADP55641 Human PRO
1975 6 5.1 315 6 ADA55373 Human pro
1976 6 5.1 315 6 AAE31485 Human but
1977 6 5.1 315 8 ADL67132 Human B7-
1978 6 5.1 316 8 ADK14172 Pneumocys
1979 6 5.1 316 8 ADK14173 Pneumocys
1980 6 5.1 316 8 ADS28065 Bacterial
1981 6 5.1 317 7 ADC87615 Human GPC
1982 6 5.1 317 8 ADX73775 Plant ful
1983 6 5.1 317 9 AEA35116 S. coccoc
1984 6 5.1 318 3 ABA23250 Streptomy
1985 6 5.1 318 7 ADC31072 Human nov
1986 6 5.1 319 7 ABO74815 Pseudomon
1987 6 5.1 320 4 AAG91867 Corynebac
1988 6 5.1 320 4 AAG91867 C glutami
1989 6 5.1 320 6 ABU22462 Protein e
1990 6 5.1 321 4 ABG03892 Novel hum
1991 6 5.1 321 6 ABU02202 S. pneumo
1992 6 5.1 321 8 ADX46605 Streptoco
1993 6 5.1 321 8 ADX77512 Plant ful
1994 6 5.1 321 9 ABM93944 M. xanthu
1995 6 5.1 322 4 ABB64332 Drosophil

1996 6 5.1 322 7 ADM25679
1997 6 5.1 322 8 ADR94420
1998 6 5.1 322 9 AEA58290
1999 6 5.1 323 4 ABB64328
2000 6 5.1 323 5 ABP74095

Adm25679 Hyperther
Adi94420 Novel S.
Aea58290 Streptoco
Abb64328 Drosophil
Abp74095 Candida a

ALIGNMENTS

RESULT 1

ID AAR63288 standard; protein; 133 AA.

XX AAR63288;

XX 25-MAR-2003 (revised)

DT 01-AUG-1995 (first entry)

XX Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
classification; immunisation; prophylaxis; serotyping.

OS Hepatitis C virus type 3.

XX WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

DR N-PSDB; AAQ78040.

XX New polynucleotide sequences from hepatitis C virus - and related
vectors, polypeptide(s) and antibodies, useful for immunisation,
treatment, diagnosis and typing of HCV isolates.

PS Claim 11; Page 125; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
nucleotides selected from an HCV type 3 genomic sequence, more
particularly (i) the region spanning positions 417-957 of the Core/E1
region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
amplify nucleic acid from an isolate belonging to a specific genotype, or
as a probe for specific detection/classification of nucleic acid.
XX Polypeptides encoded by the nucleotides in such compositions may be used
for immunisation against HCV, for the detection of antibodies directed
against HCV and for serotyping. This sequence corresponds to the NS3/NS4
region of HCV subtype 3a and is taken from a clone designated BR36-20-
164. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 133 AA;

Query Match 100.0%; Score 118; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 2.4e-103;

Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMGADLEVTTSWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 60

Db 16 ACMGADLEVTTSWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 75

Qy 61 EMECSQAAPYIEQAQVIAHQFKGKVLGLLQRAQQQAVIEPIVTTNWKLEAFWHKH 118
Db 76 EMECSQAAPYIEQAQVIAHQFKGKVLGLLQRAQQQAVIEPIVTTNWKLEAFWHKH 133

RESULT 2

AAR37937

ID AAR37937 standard; protein; 128 AA.

XX AAR37937;

DT 25-MAR-2003 (revised)

DT 23-SEP-1993 (first entry)

XX HCV NS4 region consensus sequence.

XX Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;

KW NS-5; phylogeny; differentiation; NS-3; core region; type 3.

XX Synthetic.

OS WO9310239-A2.

XX 27-MAY-1993.

XX 20-NOV-1992; 92WO-GB002143.

XX 21-NOV-1991; 91GB-00024596.

PR 24-JUN-1992; 92GB-00013362.

XX (COMM-) COMMON SERVICES AGENCY.

XX Simmonds P, Chan S, Yap PL;

XX WPI; 1993-182554/22.

DR N-PSDB; AAQ43111.

XX DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
diagnosing and treating HCV infection, screening blood samples and
identifying different HCV types.

PS Disclosure; Fig 9b; 120pp; English.

XX The sequences given in AAR37932-37 show amino acids 1638-1765 of the NS4
region of hepatitis C virus-3 (HCV-3) samples from 5 blood donors and a
consensus sequence. Analysis of this and other regions of the HCV genome
revealed the existence of three distinct groups of HCV. Analysis of the
region encompassing -255 to -62 of the 5' non coding region (see AAQ43058
-75) showed a difference of 9-14% in the nucleotide sequences between the
three groups. Two of the groups identified were similar to those of HCV
variants termed type 1 and 2, whilst the third appeared to represent a
novel type of virus. Comparison of the NS3 region (see AAR37927-30)
showed a high degree of sequence diversity with type 3 being
phylogenetically different to type 1 and 2. The same degree
of differentiation was noted in the NS-5 (see AAR37923-26) and core region
between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct
PN field.)

XX Sequence 128 AA;

Query Match 70.3%; Score 83; DB 2; Length 128;

Best Local Similarity 100.0%; Pred. No. 2.9e-70;

Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMGADLEVTTSWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 60

Db 10 ACMGADLEVTTSWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPAIVDPKEVLYQQYD 69

Qy 61 EMECSQAAPYIEQAQVIAHQFK 83

Db 70 EMECSQAAPYIEQAQVIAHQFK 92

```
RESULT 3
AAR37932
ID AAR37932 standard; protein; 128 AA.
XX
AC AAR37932;
XX
XX
XX 25-MAR-2003 (revised)
DT 23-SEP-1993 (first entry)
XX
XX HCV NS4 region from donor T0040.
XX
XX Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
KW NS-5; phylogeny; differentiation; NS-3; core region; type 3.
XX
XX Hepatitis C virus.
XX
XX WO9310239-A2.
XX
XX 27-MAY-1993.
XX
XX 20-NOV-1992; 92WO-GB002143.
XX
XX 21-NOV-1991; 91GB-00024696.
PR 24-JUN-1992; 92GB-00013362.
XX
XX (COMM-) COMMON SERVICES AGENCY.
XX
XX Simmonds P, Chan S, Yap PL;
PI
XX WPI; 1993-182554/22.
DR N-PSDB; AAQ43106.
XX
XX DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
PT diagnosing and treating HCV infection, screening blood samples and
PT identifying different HCV types.
XX
XX Disclosure; Fig 9b; 120pp; English.
XX
XX The sequences given in AAR37932-37 show amino acids 1638-1765 of the NS4
CC region of hepatitis C virus-3 (HCV-3) samples from 5 blood donors and a
CC consensus sequence. Analysis of this and other regions of the HCV genome
CC revealed the existence of three distinct groups of HCV. Analysis of the
CC region encompassing -255 to -62 of the 5' non coding region (see AAQ43059
CC -75) showed a difference of 9-14% in the nucleotide sequences between the
CC three groups. Two of the groups identified were similar to those of HCV
CC variants termed type 1 and 2, whilst the third appeared to represent a
CC novel type of virus. Comparison of the NS3 region (see AAR37927-30)
CC showed a high degree of sequence diversity with type 3 being
CC phylogenetically different to type 1 and 2. The same degree
CC differentiation was noted in the NS-5 (see AAR37923-26) and core region
CC between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
XX Sequence 128 AA;
SQ
Query Match 70.3%; Score 83; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.9e-70;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSGVCVVIHGHIELGGKPAIVDPKVELVYQYD 60
Db 10 ACMSADLEVTSTWVLLGGVLAALAAAYCLSGVCVVIHGHIELGGKPAIVDPKVELVYQYD 69
Qy 61 EMECSQAAPYIEQAQVIAHQFK 83
Db 70 EMECSQAAPYIEQAQVIAHQFK 92
RESULT 4
AAR63289
ID AAR63289 standard; protein; 133 AA.
XX
XX
XX
XX 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX
XX Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
XX
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW classification; immunisation; prophylaxis; serotyping.
XX
XX Hepatitis C virus type 3.
XX
XX WO9425601-A2.
XX
XX 10-NOV-1994.
XX
XX 27-APR-1994; 94WO-EP001323.
XX
XX 27-APR-1993; 93EP-00401099.
PR 05-AUG-1993; 93EP-00402019.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
XX Maertens G, Stuyver L;
PI
XX WPI; 1994-358277/44.
DR N-PSDB; AAQ78041.
XX
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX
XX Claim 11; Page 127; 404pp; English.
XX
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a, or (v) an HCV subtype 3c genomic sequence, may be used as primers to
CC amplify nucleic acid from an isolate belonging to a specific genotype, or
CC as a probe for specific detection/classification of nucleic acid.
CC Polypeptides encoded by the nucleotides in such compositions may be used
CC for immunisation against HCV, for the detection of antibodies directed
CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
CC region of HCV subtype 3a and is taken from a clone designated BR36-20-
CC 166. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 133 AA;
SQ
Query Match 70.3%; Score 83; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 3e-70;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSGVCVVIHGHIELGGKPAIVDPKVELVYQYD 60
Db 16 ACMSADLEVTSTWVLLGGVLAALAAAYCLSGVCVVIHGHIELGGKPAIVDPKVELVYQYD 75
Qy 61 EMECSQAAPYIEQAQVIAHQFK 83
Db 76 EMECSQAAPYIEQAQVIAHQFK 98
RESULT 5
AAR63290
ID AAR63290 standard; protein; 133 AA.
XX
XX
XX
XX 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX
```

XX Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
 XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping.
 XX Hepatitis C virus type 3.
 XX WO9425601-A2.
 XX 10-NOV-1994.
 XX 27-APR-1994; 94WO-EP001323.
 XX 27-APR-1993; 93EP-00401099.
 PR 05-AUG-1993; 93EP-00402019.
 XX (INNO-) INNOGENETICS NV SA.
 PA Maertens G, Stuyver L;
 FI WPI; 1994-358277/44.
 XX N-PSDB; AAQ78042.
 XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates.
 XX Claim 11; Page 128-129; 404pp; English.
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or
 CC as a probe for specific detection/classification of nucleic acid.
 CC Polypeptides encoded by the nucleotides in such compositions may be used
 CC for immunisation against HCV, for the detection of antibodies directed
 CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
 CC region of HCV subtype 3a and is taken from a clone designated BR36-20-
 CC 165. (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 133 AA;
 SQ
 Query Match 70.3%; Score 83; DB 2; Length 133;
 Best Local Similarity 100.0%; Pred. No. 3e-70;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKEVLYQQYD 60
 Db 16 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKEVLYQQYD 75
 Qy 61 EMECSQAAPYIEQAQVIAHQFK 83
 Db 76 EMECSQAAPYIEQAQVIAHQFK 98
 RESULT 6
 AAR63390
 ID AAR63390 standard; protein; 209 AA.
 XX AAR63390;
 AC AAR63390;
 XX 25-MAR-2003 (revised)
 DT 18-AUG-1995 (first entry)
 XX HCV polypeptide sequence.
 DE Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW

KW classification; immunisation; prophylaxis; serotyping.
 XX Hepatitis C virus.
 XX WO9425601-A2.
 XX 10-NOV-1994.
 XX 27-APR-1994; 94WO-EP001323.
 XX 27-APR-1993; 93EP-00401099.
 PR 05-AUG-1993; 93EP-00402019.
 XX (INNO-) INNOGENETICS NV SA.
 PA Maertens G, Stuyver L;
 FI WPI; 1994-358277/44.
 XX N-PSDB; AAQ78125.
 XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates.
 XX Disclosure; Page 274-275; 404pp; English.
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or
 CC as a probe for specific detection/classification of nucleic acid.
 CC Polypeptides encoded by the nucleotides in such compositions may be used
 CC for immunisation against HCV, for the detection of antibodies directed
 CC against HCV and for serotyping. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX Sequence 209 AA;
 SQ
 Query Match 70.3%; Score 83; DB 2; Length 209;
 Best Local Similarity 100.0%; Pred. No. 4.3e-70;
 Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKEVLYQQYD 60
 Db 92 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGVVIVGHIELGKPAIVDPKEVLYQQYD 151
 Qy 61 EMECSQAAPYIEQAQVIAHQFK 83
 Db 152 EMECSQAAPYIEQAQVIAHQFK 174
 RESULT 7
 AAR37934
 ID AAR37934 standard; protein; 117 AA.
 XX AAR37934;
 AC AAR37934;
 XX 25-MAR-2003 (revised)
 DT 23-SEP-1993 (first entry)
 XX HCV NS4 region from donor T0036.
 DE Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
 KW NS-5; phylogeny; differentiation; NS-3; core region; type 3.
 XX Hepatitis C virus.
 OS
 XX WO9310239-A2.
 PN


```
XX SQ Sequence 133 AA;
Query Match 39.8%; Score 47; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.2e-36;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 47
Db 16 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 62

RESULT 12
AAR63286
ID AAR63286 standard; protein; 133 AA.
XX
AC AAR63286;
XX
DT 25-MAR-2003 (revised)
DT 01-AUG-1995 (first entry)
XX
DE Polypeptide encoded by hepatitis C virus NS3/NS4 sequence.
XX
KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
KW Classification; immunisation; prophylaxis; serotyping.
XX
OS Hepatitis C virus type 3.
XX
PN WO9425601-A2.
XX
PD 10-NOV-1994.
XX
PF 27-APR-1994; 94WO-EP001323.
XX
PR 27-APR-1993; 93EP-00401099.
PR 05-AUG-1993; 93EP-00402019.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI Maertens G, Stuyver L;
XX
DR WPI; 1994-358277/44.
DR N-PSDB; AAQ78036.
XX
New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX
Claim 11; Page 121-122; 404pp; English.
XX
Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the PR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
CC amplify nucleic acid from an isolate belonging to a specific genotype, or
CC as a probe for specific detection/classification of nucleic acid.
CC Polypeptides encoded by the nucleotides in such compositions may be used
CC for immunisation against HCV, for the detection of antibodies directed
CC against HCV and for serotyping. This sequence corresponds to the NS3/NS4
CC region of HCV subtype 3a and is taken from a clone designated HD10-1-25.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 133 AA;
Query Match 39.8%; Score 47; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.2e-36;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 47
Db 16 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 62

RESULT 13
ADK40823
ID ADK40823 standard; protein; 3021 AA.
XX
AC ADK40823;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #46.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
SQ Sequence 3021 AA;
Query Match 39.8%; Score 47; DB 9; Length 3021;
Best Local Similarity 100.0%; Pred. No. 4.4e-35;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 47
Db 1653 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 1699

RESULT 14
ADK40824
ID ADK40824 standard; protein; 3021 AA.
XX
AC ADK40824;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #47.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
```

PN WO2005012502-A2.
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3021 AA;
XX
XX Query Match 35.6%; Score 42; DB 9; Length 3021;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-30;
XX Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHGKPA 47
XX 1653 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHGKPA 1699
XX
XX RESULT 15
XX ID ADX40820 standard; protein; 3021 AA.
XX
XX AC ADX40820;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE HCV polymerase protein #43.
XX
XX KW Immune stimulation; polymerase; enzyme.
XX
XX OS Hepatitis C virus.
XX
XX PN WO2005012502-A2.
XX
XX PD 10-FEB-2005.
XX
XX PF 29-MAR-2004; 2004WO-US009510.
XX
XX PR 28-MAR-2003; 2003US-0458026P.
XX
XX PA (EPIM-) EPIMMUNE INC.
XX
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX DR WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
PT

XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3021 AA;
XX
XX Query Match 35.6%; Score 42; DB 9; Length 3021;
XX Best Local Similarity 100.0%; Pred. No. 2.3e-30;
XX Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHGKPA 42
XX 1653 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVVIGHGKPA 1694
XX
XX RESULT 16
XX ID AAE18690 standard; protein; 829 AA.
XX
XX AC AAE18690;
XX
XX DT 17-MAY-2002 (first entry)
XX
XX DE Multiple epitope fusion antigen (MEFA) 12 protein.
XX
XX KW Hepatitis C virus; NS3/4a antigen; multiple epitope fusion antigen;
XX HCV infection; MEFA 12 protein.
XX
XX OS Unidentified.
XX
XX PH Key Location/Qualifiers
XX FT Misc-difference 315 /note= "Encoded by ATG"
XX FT Misc-difference 645 /note= "Encoded by GAG"
XX
XX PN WO200196875-A2.
XX
XX PD 20-DEC-2001.
XX
XX PF 14-JUN-2001; 2001WO-US019369.
XX
XX PR 15-JUN-2000; 2000US-0212082P.
XX
XX PR 02-APR-2001; 2001US-0280811P.
XX
XX PR 02-APR-2001; 2001US-0280867P.
XX
XX PA (CHIR) CHIRON CORP.
XX
XX PI Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
XX Medina-Selby A;
XX
XX DR WPI; 2002-179522/23.
XX
XX DR N-PSDB; AAD29796.
XX
XX PT Immunoassay solid support useful for detecting hepatitis C virus
PT infection in a biological sample, comprises at least one of HCV anti-core
PT antibody and HCV NS3/4a epitope, bound to the support.
XX
XX PS Disclosure; Fig 7; 87pp; English.
XX
XX The present invention relates to hepatitis C virus (HCV) core antigen and
CC NS (nonstructural) 3/4a antibody combination assay that can detect both
CC HCV antigens and antibodies present in a sample using a single solid
CC matrix as well as immunoassay solid supports for use in the assay. The
CC solid support is useful for detecting HCV infection in a biological

CC sample. The present sequence is MEPA (multiple epitope fusion antigen) 12
CC protein. This sequence is used in the exemplification of the invention
SQ Sequence 829 AA;

Query Match 34.7%; Score 41; DB 5; Length 829;
Best Local Similarity 100.0%; Pred. No. 7e-30;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 43 GGPALVDPKVELVYQQYDEMEECSQAAPYIEQAQVIAHQFK 83
Db 444 GGPALVDPKVELVYQQYDEMEECSQAAPYIEQAQVIAHQFK 484

RESULT 17
ADCO6769
ID ADCO6769 standard; protein; 829 AA.

XX AC ADCO6769;

XX DT 18-DEC-2003 (first entry)

XX DE Chimeric multiple epitope fusion antigen 12 protein.

XX KW immunoassay solid support; HCV; NS3/4a; non-structural;
XX KW non-A, non-B hepatitis; NANB; multiple epitope fusion antigen 12; MEPA12;
XX KW chimeric.

XX OS Chimeric.

XX OS Synthetic.

XX OS Unidentified.

XX OS Hepatitis C virus.

XX OS Homo sapiens.

XX XX US2002192639-A1.

XX XX 19-DEC-2002.

XX XX 14-JUN-2001; 2001US-00881239.
XX XX 15-JUN-2000; 2000US-0212082P.

XX PR 02-APR-2001; 2001US-0280811P.

XX PR 02-APR-2001; 2001US-0280867P.

XX XX (CHIE)/ CHIEN D Y.

XX PA (ARCA)/ ARCANDEL P.

XX PA (TAND)/ TANDESKE L.

XX PA (GEOR)/ GEORGE-NASCIMENTO C.

XX PA (COIT)/ COIT D.

XX PA (MEDI)/ MEDINA-SELBY A.

XX PI Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
XX PI Medina-Selby A;

XX DR WPI; 2003-644609/61.

XX DR N-PSDB; ADCO6770.

XX XX Claim 45; Fig 7; 40pp; English.

XX The invention relates to a novel immunoassay solid support comprising at
XX least one hepatitis C virus (HCV) anti-core antibody and at least one
XX isolated HCV NS3/4a (non-structural protein 3/4a) epitope bound thereto.
XX The system of the invention may be useful for detecting HCV infection in
XX a biological sample and for treating or detecting non-A, non-B hepatitis
XX (NANB hepatitis). The current sequence is that of the chimeric multiple
XX epitope fusion antigen 12 (MEPA12) protein of the invention.

XX XX Sequence 829 AA;

XX Query Match 34.7%; Score 41; DB 8; Length 829;
XX Best Local Similarity 100.0%; Pred. No. 7e-30;
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 43 GGPALVDPKVELVYQQYDEMEECSQAAPYIEQAQVIAHQFK 83

Query Match 34.7%; Score 41; DB 7; Length 829;
Best Local Similarity 100.0%; Pred. No. 7e-30;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GGPALVDPKVELVYQQYDEMEECSQAAPYIEQAQVIAHQFK 83
Db 444 GGPALVDPKVELVYQQYDEMEECSQAAPYIEQAQVIAHQFK 484

RESULT 18

ADL66807

ID ADL66807 standard; protein; 829 AA.

XX AC ADL66807;

XX DT 03-JUN-2004 (first entry)

XX DE HCV multiple epitope fusion antigen 12 (MEPA 12) polypeptide.

XX KW HCV; MEPA 12; HCV antigen; HCV polypeptide;

XX KW multiple epitope fusion antigen; MEPA; hepatitis C virus infection;

XX KW multiple epitope fusion antigen 12.

XX OS Hepatitis C virus.

XX XX WO2004021871-A2.

XX PD 18-MAR-2004.

XX PF 08-SEP-2003; 2003WO-US028071.

XX PR 09-SEP-2002; 2002US-0409515P.

XX XX (CHIR) CHIRON CORP.

XX PI Arcangel P, Chien D;

XX DR WPI; 2004-248333/23.

XX DR N-PSDB; ADL66806.

XX The invention relates to a method of detecting hepatitis C virus (HCV)
XX infection in a biological sample. The method comprises providing an
XX immunoassay solid support comprising HCV antigens bound to it, where the
XX HCV antigens comprise one or more isolated antigens from a first region
XX of the HCV polypeptide, combining a biological sample with the solid
XX support under conditions that allow HCV antibodies, when present in the
XX biological sample, to bind to the one or more HCV antigens, adding to the
XX solid support a detectably labelled HCV multiple epitope fusion antigen
XX (MEPA), where the labelled MEPA comprises at least one epitope from the
XX same region of the HCV polypeptide as the one or more isolated antigens,
XX where the MEPA binds to the bound HCV antibody, and detecting complexes
XX formed between the HCV antibody and the one or more antigens from the
XX first region of the HCV polypeptide and the MEPA, if any, as an
XX indication of HCV infection in the biological sample. The method is
XX useful for detecting hepatitis C virus (HCV) infection in a biological
XX sample. This sequence represents the MEPA 12 polypeptide used in the
XX scope of the invention.

XX PS Claim 14; SEQ ID NO 4; 93pp; English.

XX The invention relates to a method of detecting hepatitis C virus (HCV)
XX infection in a biological sample. The method comprises providing an
XX immunoassay solid support comprising HCV antigens bound to it, where the
XX HCV antigens comprise one or more isolated antigens from a first region
XX of the HCV polypeptide, combining a biological sample with the solid
XX support under conditions that allow HCV antibodies, when present in the
XX biological sample, to bind to the one or more HCV antigens, adding to the
XX solid support a detectably labelled HCV multiple epitope fusion antigen
XX (MEPA), where the labelled MEPA comprises at least one epitope from the
XX same region of the HCV polypeptide as the one or more isolated antigens,
XX where the MEPA binds to the bound HCV antibody, and detecting complexes
XX formed between the HCV antibody and the one or more antigens from the
XX first region of the HCV polypeptide and the MEPA, if any, as an
XX indication of HCV infection in the biological sample. The method is
XX useful for detecting hepatitis C virus (HCV) infection in a biological
XX sample. This sequence represents the MEPA 12 polypeptide used in the
XX scope of the invention.

XX SQ Sequence 829 AA;

Query Match 34.7%; Score 41; DB 8; Length 829;
Best Local Similarity 100.0%; Pred. No. 7e-30;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GGPALVDPKVELVYQQYDEMEECSQAAPYIEQAQVIAHQFK 83

Db 444 GGGKPAIVDPKEVLYQQYDMEECSSQAAPYIEQAQVIAHQFK 484
|||||
AAU76378
ID AAU76378 standard; protein; 1099 AA.
AC AAU76378;
XX
XX
DT 08-MAY-2002 (first entry)
XX
XX HCV multiple epitope fusion antigen (MEFA) 7.1 protein sequence.
XX
XX Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
KW immunoassay solid support; multiple epitope fusion antigen; MEFA;
KW non-structural protein.
XX
XX Hepatitis C virus.
OS Synthetic.
XX
XX WO200196870-A2.
FN
PD 20-DEC-2001.
XX
XX 14-JUN-2001; 2001US-0210822P.
XX
XX 15-JUN-2001; 2001US-0210822P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
XX (CHIR) CHIRON CORP.
PA
XX Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;
XX
XX WPI; 2002-090228/12.
DR N-PSDB; ASK15345.
XX
XX Immunoassay solid support, useful for detecting hepatitis C virus
PT infection in biological sample, comprises HCV NS3/4a conformational
FT epitope and multiple epitope fusion antigen bound to the support.
XX
XX Claim 5; Fig 5; 92pp; English.
PS
XX The present invention relates to a new immunoassay solid support
CC consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
CC conformational epitope and a multiple epitope fusion antigen (MEFA),
CC bound to the support. The NS3/4a conformational epitope and/or MEFA,
CC reacts specifically with anti-HCV antibodies present in a biological
CC sample from an HCV-infected individual. The immunoassay of the invention
CC is useful for detecting hepatitis C virus infection in a biological
CC sample. The method of the invention provides a sensitive, accurate
CC diagnostic and prognostic tool to provide adequate patient care and to
CC prevent transmission of HCV by blood and by blood products, or by
CC personal contact. Use of NS3/4a conformational epitope in combination
CC with MEFA, provides a sensitive and reliable method for detecting early
CC HCV seroconversion. Use of MEFA has the added advantages of decreasing
CC masking problems, improving sensitivity in detecting antibodies by
CC allowing a greater number of epitopes on a unit surface area of
CC substrate, and improving substrate. Detection accuracy is increased and
CC the incidence of false results is reduced because of the identification
CC and the use of highly immunogenic HCV antigens which are present during
CC the early stages of HCV seroconversion. The present amino acid sequence
CC represents the multiple epitope fusion antigen (MEFA) 7.1 of the
XX invention
XX
XX Sequence 1099 AA;
SQ
Query Match 34.7%; Score 41; DB 5; Length 1099;
Best Local Similarity 100.0%; Pred. No. 8.8e-30;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX US2002146685-A1.
XX 10-OCT-2002.
XX

QY 43 GGGKPAIVDPKEVLYQQYDMEECSSQAAPYIEQAQVIAHQFK 83
|||||
Db 748 GGGKPAIVDPKEVLYQQYDMEECSSQAAPYIEQAQVIAHQFK 788
|||||
RESULT 20
ABG72262
ID ABG72262 standard; protein; 1099 AA.
XX
XX AC ABG72262;
XX
XX DT 06-MAR-2003 (first entry)
XX
XX HCV multiple epitope fusion antigen 7.1 (MEFA 7.1).
DE
XX Immunoassay solid support; Hepatitis C Virus type-1; HCV-1; HCV-2;
KW NS3/4a conformational epitope; multiple epitope fusion antigen 7.1;
KW MEFA 7.1; anti-HCV antibody; NS3/4a conformational antigen; HCV-3;
KW HCV infection; Hepatitis C Virus type-2; Hepatitis C Virus type-3;
KW mutant; mutein.
XX
XX Hepatitis C virus type 1.
OS Hepatitis C virus type 2.
OS Hepatitis C virus type 3.
OS Synthetic.
OS Chimeric.
XX
XX Key
PH Location/Qualifiers
FT 1..156
FT /note= "Correspond to amino acids 1-156 of HCV-1 hSOD
FT superoxide dismutase"
FT 159..176
FT /note= "Correspond to amino acids 303-320 of HCV-1 E1"
FT 179..199
FT /note= "Correspond to consensus sequence of amino acids
FT 390-410 of HCV-1 E2 HVR"
FT 200..230
FT /note= "Correspond to consensus sequence of amino acids
FT 384-414 of HCV-1 and HCV-2 E2 HVR"
FT 231..696
FT /note= "Correspond to amino acids 1193-1658 of HCV-1
FT helicase"
FT 699..745
FT /note= "Correspond to amino acids 1689-1735 of HCV-1 5-1-
FT 1 epitope"
FT 748..794
FT /note= "Correspond to amino acids 1689-1735 of HCV-3 5-1-
FT 1 epitope"
FT 797..843
FT /note= "Correspond to amino acids 1689-1735 of HCV-2 5-1-
FT 1 epitope"
FT 846..881
FT /note= "Correspond to amino acids 1901-1936 of HCV-1
FT polypeptide C100"
FT 884..919
FT /note= "Correspond to amino acids 2278-2313 of HCV-1 NS5
FT region"
FT 922..957
FT /note= "Correspond to amino acids 2278-2313 of HCV-1 NS5
FT region"
FT 958..1028
FT /note= "Correspond to core region antigenic determinants
FT from amino acids 9-32, 39-42 and 64-88 of HCV-1 and amino
FT acids 67-84 of HCV-2"
FT 1029..1099
FT /note= "Correspond to core region antigenic determinants
FT from amino acids 9-32, 39-42 and 64-88 of HCV-1 and amino
FT acids 67-84 of HCV-2"
XX
XX US2002146685-A1.
XX 10-OCT-2002.
XX

```

PF 14-JUN-2001; 2001US-00881554.
XX
PR 15-JUN-2000; 2000US-0212082P.
PR 02-APR-2001; 2001US-0280811P.
PR 02-APR-2001; 2001US-0280867P.
XX
XX (CHIE/) CHIEN D Y.
PA (ARCA/) ARCANDEL P.
PA (TAND/) TANDESKE L.
PA (GEOR/) GEORGE-NASCIMENTO C.
PA (COIT/) COIT D.
PA (MEDI/) MEDINA-SELBY A.
XX
XX Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;
XX
XX WPI; 2003-147573/14.
XX N-PSDB; ABX14411.
XX
XX Immunoassay solid support for detecting Hepatitis C Virus infection in
PT biological samples, comprises Hepatitis C Virus conformational epitope
PT and multiple epitope fusion antigen.
XX
XX Claim 25; Fig 5A-5F; 45pp; English.
XX
XX The present invention relates to immunoassays comprising Hepatitis C
XX Virus (HCV) NS3/4a conformational epitope and multiple epitope fusion
XX antigen (MEFA), bound to a solid support. The NS3/4a epitope and/or the
XX multiple epitope fusion antigen react with anti-HCV antibodies present in
XX a biological sample from an HCV-infected individual. The immunoassays and
XX methods of the invention are useful for detecting HCV infection in a
XX biological sample. The inventive immunoassay solid support provides a
XX sensitive and reliable method for detecting early HCV seroconversion. The
XX assays can detect HCV infection caused by any six known genotypes of HCV.
XX The use of the multiple epitope fusion proteins decreases masking
XX problems, improves sensitivity in detecting antibodies by allowing a
XX greater number of epitopes on a unit area of substrate, and improves
XX selectivity. The present sequence represents HCV multiple epitope fusion
XX antigen 7.1 (MEFA 7.1), a mutant HCV polypeptide derived from various
XX regions of HCV type 1, 2, or 3 (HCV-1, HCV-2, or HCV-3) polypeptide
XX sequences
XX
XX Sequence 1099 AA;
XX
XX Query Match 34.7%; Score 41; DB 6; Length 1099;
XX Best Local Similarity 100.0%; Pred. No. 8.8e-30;
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 43 GGRPAIVDPKVELVYQQYDEMEECSSQAAPYIEQAQVIAHQFK 83
XX 748 GGRPAIVDPKVELVYQQYDEMEECSSQAAPYIEQAQVIAHQFK 788
XX
XX RESULT 21
XX ADL66809
XX ID ADL66809 standard; protein; 1099 AA.
XX
XX AC ADL66809;
XX
XX DT 03-JUN-2004 (first entry)
XX
XX HCV multiple epitope fusion antigen 7.1 (MEFA 7.1) polypeptide.
XX
XX HCV; MEFA 7.1; HCV antigen; HCV polypeptide;
XX multiple epitope fusion antigen; MEFA; hepatitis C virus infection;
XX multiple epitope fusion antigen 7.1.
XX
XX Hepatitis C virus.
XX
XX WO2004021871-A2.
XX
XX 18-MAR-2004.
XX
PF 08-SEP-2003; 2003WO-US028071.
XX
PR 09-SEP-2002; 2002US-0409515P.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Arcangel P, Chien D;
XX
XX WPI; 2004-248333/23.
XX N-PSDB; ADL66808.
XX
XX Detecting hepatitis C virus (HCV) infection in a biological sample by
XX detecting complexes formed between the HCV antibody and the antigens from
XX the first region of the HCV polypeptide and the multiple epitope fusion
XX antigen (MEFA).
XX
XX Claim 15; SEQ ID NO 6; 93pp; English.
XX
XX The invention relates to a method of detecting hepatitis C virus (HCV)
XX infection in a biological sample. The method comprises providing an
XX immunoassay solid support comprising HCV antigens bound to it, where the
XX HCV antigens comprise one or more isolated antigens form a first region
XX of the HCV polypeptide, combining a biological sample with the solid
XX support under conditions that allow HCV antibodies, when present in the
XX biological sample, to bind to the one or more HCV antigens, adding to the
XX solid support a detectably labelled HCV multiple epitope fusion antigen
XX (MEFA), where the labelled MEFA comprises at least one epitope from the
XX same region of the HCV polypeptide as the one or more isolated antigens,
XX where the MEFA binds to the bound HCV antibody, and detecting complexes
XX formed between the HCV antibody and the one or more antigens from the
XX first region of the HCV polypeptide and the MEFA, if any, as an
XX indication of HCV infection in the biological sample. The method is
XX useful for detecting hepatitis C virus (HCV) infection in a biological
XX sample. This sequence represents the MEFA 7.1 polypeptide used in the
XX scope of the invention.
XX
XX Sequence 1099 AA;
XX
XX Query Match 34.7%; Score 41; DB 8; Length 1099;
XX Best Local Similarity 100.0%; Pred. No. 8.8e-30;
XX Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 43 GGRPAIVDPKVELVYQQYDEMEECSSQAAPYIEQAQVIAHQFK 83
XX 748 GGRPAIVDPKVELVYQQYDEMEECSSQAAPYIEQAQVIAHQFK 788
XX
XX RESULT 22
XX AAR94462
XX ID AAR94462 standard; protein; 3023 AA.
XX
XX AC AAR94462;
XX
XX DT 20-SEP-1996 (first entry)
XX
XX Hepatitis C virus polypeptide.
XX
XX Hepatitis C virus; antibody; detection; diagnosis; vaccine; classify;
XX subtype.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX Peptide 1505..1520
XX /note="this part of the sequence is missing from the
XX specification"
XX
XX Peptide 2433..2448
XX /note="this part of the sequence is missing from the
XX specification"
XX
XX JP08056672-A.
XX
XX 05-MAR-1996.

```



```

XX AC AAR63316;
XX
XX DT 16-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 09-AUG-1995 (first entry)
XX
XX DE Peptide fragment of hepatitis C virus type 3.
XX
XX KW Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
XX KW classification; immunisation; prophylaxis; serotyping.
XX
XX OS Hepatitis C virus; isolate BR36.
XX
XX PN WO9425601-A2.
XX
XX PD 10-NOV-1994.
XX
XX PF 27-APR-1994; 94WO-EP001323.
XX
XX PR 27-APR-1993; 93EP-00401099.
XX PR 05-AUG-1993; 93EP-00402019.
XX
XX PA (INNO-) INNOGENETICS NV SA.
XX
XX PI Maertens G, Stuyver L;
XX
XX DR WPI; 1994-358277/44.
XX
XX PT New polynucleotide sequences from hepatitis C virus - and related
XX PT vectors, polypeptide(s) and antibodies, useful for immunisation,
XX PT treatment, diagnosis and typing of HCV isolates.
XX
XX PS Disclosure; Page 166; 404pp; English.
XX
XX CC Compositions comprising at least 5, and pref. 8 or more contiguous
XX CC nucleotides selected from an HCV type 3 genomic sequence, more
XX CC particularly (i) the region spanning positions 417-957 of the Core/E1
XX CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
XX CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
XX CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
XX CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
XX CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
XX CC genomic sequence, a type 4 genomic sequence; or the coding region of
XX CC subtype 5a, may be used as primers to amplify nucleic acid from an
XX CC isolate belonging to a specific genotype, or as a probe for specific
XX CC detection/classification of nucleic acid. Polypeptides encoded by the
XX CC nucleotides in such compositions may be used for immunisation against
XX CC HCV, for the detection of antibodies directed against HCV and for
XX CC serotyping. This polypeptide corresponds to positions 1688-1707 of HCV
XX CC type 3. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-
XX CC 2003 to standardise OS field)
XX
XX SQ Sequence 20 AA;
XX
XX Query Match 16.9%; Score 20; DB 2; Length 20;
XX Best Local Similarity 100.0%; Pred. No. 2.2e-11;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 42 LGGKPAIVDPKENVLYQQYDE 61
XX | | | | | | | | | | | | | | | | | |
XX Db 1 LGGKPAIVDPKENVLYQQYDE 20
XX
XX RESULT 26
XX AAR41164
XX ID AAR41164 standard; peptide; 22 AA.
XX AC AAR41164;
XX
XX DT 25-MAR-2003 (revised)
XX DT 22-MAR-1994 (first entry)
XX

```

```

DE XX HCV (type 3) peptide NS4-1 (3).
XX
XX KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
XX KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
XX KW epitope; antibody; biotin; diagnosis; detection; vaccine.
XX
XX OS Synthetic.
XX
XX PH Key Location/Qualifiers
XX FT Modified-site 1
XX FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B=
XX FT biotin; X= biotinylation cpd. incorporated during
XX FT synthesis; Y= bond or linking gp(s). which minimises
XX FT steric hindrance, where Y is not a bond it is pref. 1-10
XX FT residues of (same or different) glycine, beta-alanine, 4-
XX FT aminobutyric acid, 5-aminovaleic acid or 6-aminohexanoic
XX FT acid; parentheses around B and X indicate opt. presence in
XX FT at the specified positions but B or X must be present in
XX FT at least one of the positions shown, B interacts with the
XX FT peptide to give a cpd. with greater diagnostic
XX FT sensitivity; A (optional)= one or more amino acids, NH2
XX FT or gp. which modifies the N-terminus; Z= one or more
XX FT amino acids, OH, NH2, or a linkage involving either of
XX FT these 2 gps."
XX
XX FT Modified-site 22
XX FT /note= "the C-terminal comprises Y-(X)-Z"
XX
XX PN WO9318054-A2.
XX
XX PD 16-SEP-1993.
XX
XX PF 08-MAR-1993; 93WO-EP000517.
XX
XX PR 06-MAR-1992; 92EP-00400598.
XX
XX PA (INNO-) INNOGENETICS NV SA.
XX
XX PI De Leys R;
XX
XX DR WPI; 1993-303397/38.
XX
XX PT New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
XX PT increased antigenicity, useful in antibodies detection and vaccines
XX PT against hepatitis C, HIV and HTLV.
XX
XX PS Claim 4; Page 90-98; 133pp; English.
XX
XX CC Peptide compns. comprise at least one and pref. a combination of two,
XX CC three, four or more biotinylated peptides chosen from the sequences given
XX CC in AAR41058-K41166. The peptides represent immunologically important
XX CC regions of viral proteins and are prepd. by solid phase peptide
XX CC synthesis. The compns. are useful for the detection of antibodies to
XX CC HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct
XX CC PN field.)
XX
XX SQ Sequence 22 AA;
XX
XX Query Match 16.9%; Score 20; DB 2; Length 22;
XX Best Local Similarity 100.0%; Pred. No. 2.4e-11;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 42 LGGKPAIVDPKENVLYQQYDE 61
XX | | | | | | | | | | | | | | | | | |
XX Db 2 LGGKPAIVDPKENVLYQQYDE 21
XX
XX RESULT 27
XX AAR37941
XX ID AAR37941 standard; protein; 19 AA.
XX AC AAR37941;
XX
XX DT 25-MAR-2003 (revised)

```


XX PS Disclosure; Page 167; 404pp; English.

XX CC Compositions comprising at least 5, and pref. 8 or more contiguous

CC nucleotides selected from an HCV type 3 genomic sequence, more

CC particularly (i) the region spanning positions 417-957 of the Core/E1

CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of

CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-

CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning

CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype

CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d

CC genomic sequence, a type 4 genomic sequence; or the coding region of

CC subtype 5a, may be used as primers to amplify nucleic acid from an

CC isolate belonging to a specific genotype, or as a probe for specific

CC detection/classification of nucleic acid. Polypeptides encoded by the

CC nucleotides in such compositions may be used for immunisation against

CC HCV, for the detection of antibodies directed against HCV and for

CC serotyping. This polypeptide corresponds to positions 1712-1731 of HCV.

CC (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 20 AA;

Query Match 15.3%; Score 18; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.7e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 SQAAPYIEQAQVIAHQPK 83

DB 1 SQAAPYIEQAQVIAHQPK 18

RESULT 30

AAR41165

ID AAR41165 standard; peptide; 22 AA.

XX AC AAR41165;

XX DT 25-MAR-2003 (revised)

XX DT 22-MAR-1994 (first entry)

XX DB HCV (type 3) peptide NS4-5 (3).

XX XH Human immunodeficiency virus; HIV; hepatitis C virus; HCV;

XX XH non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;

XX XH epitope; antibody; biotin; diagnosis; detection; vaccine.

XX OS Synthetic.

XX PH Key

FT Modified-site 1

FT Location/Qualifiers

FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B=

FT biotin; X= biotinylation cpd. incorporated during

FT synthesis; Y= bond or linking gp(s). which minimises

FT steric hindrance, where Y is not a bond it is pref. 1-10

FT residues of (same or different) glycine, beta-alanine, 4-

FT aminobutyric acid, 5-aminovaleric acid or 6-aminohexanoic

FT acid; parenthesis around B and X indicate opt. presence

FT at the specified positions but B or X must be present in

FT at least one of the positions shown, B interacts with the

FT peptide to give a cpd. with greater diagnostic

FT sensitivity; A (optional)= one or more amino acids, NH2

FT or gp. which modifies the N-terminus; Z= one or more

FT amino acids, OH, NH2, or a linkage involving either of

FT these 2 gps."

FT Modified-site 22

FT /note= "the C-terminal comprises Y-(X)-Z"

XX XH W09318054-A2.

XX PD 16-SEP-1993.

XX XH 08-MAR-1993; 93WO-EP000517.

XX XH

PR 06-MAR-1992; 92EP-00400598.

XX PA (INNO-) INNOGENETICS NV SA.

XX PI De Leys R;

XX DR WPI; 1993-303397/38.

XX PT New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with

PT increased antigenicity, useful in antibodies detection and vaccines

PT against hepatitis C, HIV and HTLV.

XX Claim 4; Page 90-98; 133pp; English.

XX CC Peptide compns. comprise at least one and pref. a combination of two,

CC three, four or more biotinylated peptides chosen from the sequences given

CC in AAR41058-R41166. The peptides represent immunologically important

CC regions of viral proteins and are prepd. by solid phase peptide

CC synthesis. The compns. are useful for the detection of antibodies to

CC HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct

CC PN field.)

XX SQ Sequence 22 AA;

Query Match 15.3%; Score 18; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 SQAAPYIEQAQVIAHQPK 83

DB 2 SQAAPYIEQAQVIAHQPK 19

RESULT 31

AAY06672

ID AAY06672 standard; protein; 352 AA.

XX AC AAY06672;

XX DT 17-JUN-1999 (first entry)

XX DE Amino acid sequence of the NS4 mosaic protein.

XX XH Mosaic protein; antigenic peptide; nucleocapsid; NC; hepatitis; REAL;

XX XH restriction endonuclease assisted ligation; vaccination; NS4 antigen.

XX OS Hepatitis C virus.

XX OS Synthetic.

XX PN W09910506-A1.

XX PD 04-MAR-1999.

XX PF 21-AUG-1998; 98WO-US017385.

XX PR 25-AUG-1997; 97US-00921887.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Khudyakov YE, Fields HA;

XX DR WPI; 1999-204671/17.

XX PT New mosaic protein, comprising a plurality of homologous antigenic

PT peptides from different genotypes of a species - useful for detecting

PT hepatitis infection in an individual.

XX Example 3; Fig 18; 66pp; English.

XX CC The invention relates to a mosaic protein, comprising a plurality of

CC homologous antigenic peptides from different genotypes of a species. The

CC antigenic peptides are from nucleocapsid (NC) proteins. A method for

CC synthesising an artificial gene that encodes the mosaic protein is also

CC provided. The method is designated restriction endonuclease assisted
CC ligation (REAL). The mosaic protein and the artificial mosaic protein are
CC useful for detecting a hepatitis infection in an individual. The mosaic
CC gene and protein is also useful for vaccination against infection,
CC especially hepatitis C. The method of synthesizing the artificial gene
CC and the resulting mosaic protein improve the sensitivity, spectrum of
CC immunoreactivity, and antigen specificity of enzyme immunoassays. This
CC provides improved detection of hepatitis C virus. The present sequence
XX represents the amino acid sequence of the NS4 mosaic protein
XX
SQ Sequence 352 AA;

Query Match 15.3%; Score 18; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SQAPYIEQAQVIAHOFK 83
DB 125 SQAPYIEQAQVIAHOFK 142
|||||

RESULT 32
ADX40827
ID ADX40827 standard; protein; 3016 AA.

AC ADX40827;
DT 21-APR-2005 (first entry)
DE HCV polymerase protein #50.
KW Immune stimulation; polymerase; enzyme.
OS Hepatitis C virus.
PN WO2005012502-A2.
XX
FD 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX

PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX

PS Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX

SQ Sequence 3016 AA;

Query Match 15.3%; Score 18; DB 9; Length 3016;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLSVGCVV 35
|||||

DB 1666 GGVLAALAAAYCLSVGCVV 1683
RESULT 33
ADX40810
ID ADX40810 standard; protein; 3019 AA.

XX AC ADX40810;
XX
XX 21-APR-2005 (first entry)
XX
XX HCV polymerase protein #33.
XX Immune stimulation; polymerase; enzyme.

OS Hepatitis C virus.
XX
XX WO2005012502-A2.
XX
XX 10-FEB-2005.

XX 29-MAR-2004; 2004WO-US009510.

XX 28-MAR-2003; 2003US-0458026P.

XX (EPIM-) EPIMUNE INC.

XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

XX WPI; 2005-132661/14.

XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.

XX Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX

SQ Sequence 3019 AA;

Query Match 15.3%; Score 18; DB 9; Length 3019;

Best Local Similarity 100.0%; Pred. No. 1.1e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEVLYQQYDEMECS 66
|||||

DB 1700 VPDKEVLYQQYDEMECS 1717
|||||

RESULT 34
ADX40780

ID ADX40780 standard; protein; 3008 AA.

XX AC ADX40780;

XX 21-APR-2005 (first entry)

XX HCV polymerase protein #3.

XX Immune stimulation; polymerase; enzyme.

OS Hepatitis C virus.

XX WO2005012502-A2.

XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX PF Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PF Disclosure; Page 387-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX CC Sequence 3008 AA;
XX Query Match 12.7%; Score 15; DB 9; Length 3008;
XX Best Local Similarity 100.0%; Pred. No. 7.8e-05;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAALAAAYCLSVG 32
XX DB 1664 GGVLAALAAAYCLSVG 1678
XX RESULT 35
XX ADX40825
XX ID ADX40825 standard; protein; 3015 AA.
XX AC ADX40825;
XX XX 21-APR-2005 (first entry)
XX DT HCV polymerase protein #48.
XX DE Immune stimulation; polymerase; enzyme.
XX KW Hepatitis C virus.
XX OS WO2005012502-A2.
XX PN 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX PF Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PF Disclosure; Page 388-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX CC Sequence 3015 AA;
XX Query Match 12.7%; Score 15; DB 9; Length 3015;
XX Best Local Similarity 100.0%; Pred. No. 7.8e-05;
XX Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 22 AALAAAYCLSVGCVVI 36
XX DB 1669 AALAAAYCLSVGCVVI 1683
XX RESULT 36
XX AAR37938
XX ID AAR37938 standard; protein; 18 AA.
XX AC AAR37938;
XX XX 25-MAR-2003 (revised)
XX DT 23-SEP-1993 (first entry)
XX DE HCV NS-4 type 3 region 1 (1691-1708) peptide.
XX KW Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
XX NS-5; phylogeny; differentiation; NS-3; core region; type 3; PCR;
XX amplify; polymerase chain reaction; primer; NS4.
XX OS Synthetic.
XX PN WO9310239-A2.
XX PD 27-MAY-1993.
XX PF 20-NOV-1992; 92WO-GB002143.
XX PR 21-NOV-1991; 91GB-00024696.
XX PR 24-JUN-1992; 92GB-00013362.
XX PA (COMM-) COMMON SERVICES AGENCY.
XX PI Simmonds P, Chan S, Yap PL;
XX WPI; 1993-182554/22.
XX PF DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
PT diagnosing and treating HCV infection, screening blood samples and
PT identifying different HCV types.
XX PF Disclosure; Page 40; 120pp; English.
XX CC The sequences given in AAR37938-47 are peptides which were derived from
CC the NS-4 region of the hepatitis C virus (HCV) protein. Analysis of
CC regions of the HCV genome revealed the existence of three distinct groups
CC of HCV. Analysis of the region encompassing -255 to -62 of the 5' non
CC coding region (NCR) (see AAQ43058-75) showed a difference of 9-14% in the
CC nucleotide sequences between the three groups. Two of the groups
CC identified were similar to those of HCV variants termed type 1 and 2,
CC whilst the third appeared to represent a novel type of virus. Comparison
CC of the NS3 region (see AAR37927-30) showed a high degree of sequence
CC diversity with type 3 being phylo- genetically different to type 1 and 2.
CC The same degree of differentiation was noted in the NS-5 (see AAR37923-
CC 26), core region (see AAR37931) and the NS4 region (see AAQ43106-111)
CC between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct

CC PN field.)
SQ Sequence 18 AA; 11.9%; Score 14; DB 2; Length 18;
Query Match
Best Local Similarity 100.0%; Pred. No. 9.5e-06; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEM 62
DB 5 VPDKEVLYQQYDEM 18
|||||

RESULT 37
AAV67804
ID AAV67804 standard; peptide; 35 AA.
XX AC AAV67804;
XX
XX 23-MAR-2000 (first entry)
DT XX
XX
DE Peptide #204 for detecting hepatitis C virus infection.
XX
KW Hepatitis C virus; HCV; increased structural stability; NS4 region;
KW diagnostic antigen.
XX
OS Synthetic.
XX
XX WO9962945-A2.
XX
XX 09-DEC-1999.
XX
XX 04-JUN-1999; 99WO-US012446.
XX
XX 05-JUN-1998; 98US-00982229P.
PR 01-SEP-1998; 98US-0098705P.
PR 15-SEP-1998; 98US-0100422P.
PR 28-JAN-1999; 99WO-US001726.
XX
XX (PEPT-) PEPTIDE SOLUTIONS INC.
XX
XX Chowdhury MA, Bernstein D, Motesbocker MA;
PI WPI; 2000-086953/07.
XX
XX Improving properties of peptides for use as diagnostic antigens or for
XX preventing or treating infections.
XX
XX Claim 55; Page 73; 83pp; English.
XX
XX This is a peptide related to the immunoreactive region of the NS4 region
XX of hepatitis C virus (HCV). The peptide is useful for detecting HCV
XX infection. The invention relates to peptides derived from HCV and also
XX HIV-1 which have been modified for use as diagnostic antigens in the
XX treatment or prevention of infection. The structural stability of the
XX peptides can be increased in four different ways; through the replacement
XX of a hydrophobic amino acid with a less hydrophobic amino acid; through
XX an increase in the amount of secondary structure (i.e. alpha helix) in
XX the peptide; through the removal of a positive charge from the peptide,
XX or through the constraint of the epitopic sequence via the formation of a
XX covalent crosslink. Modified peptides of the invention are used to detect
XX infectious agents specifically HCV. Other detectable agents include HIV-1
XX Group O viruses; human T-cell lymphotropic virus-I or -II; and the
XX causative agent of syphilis. The peptides can be used for prevention or
XX treatment of infections (e.g. as vaccines, or where expressed from a
XX transgene). More generally almost any peptide can be similarly modified,
XX e.g. cytokines or interferons; major histocompatibility complex antigens;
XX hormones; growth factors; tumour markers or suppressors, or antigens from
XX many other pathogens
SQ Sequence 35 AA;
Query Match 11.9%; Score 14; DB 3; Length 35;

Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEM 62
DB 21 VPDKEVLYQQYDEM 34
|||||

RESULT 38
ADV23000
ID ADV23000 standard; peptide; 18 AA.
XX AC ADV23000;
XX
XX 10-MAR-2005 (first entry)
DT XX
XX
DE HCV H77 immunogenic peptide #241.
XX
XX Vaccine; virucide; antigen; autoimmune disease; infection;
KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
KW Hodgkin's lymphoma.
XX
OS Hepatitis C virus strain H77.
XX
XX WO2004108753-A1.
XX
XX 16-DEC-2004.
XX
XX 10-JUN-2004; 2004WO-AU000775.
XX
XX 10-JUN-2003; 2003AU-00902875.
PR 25-MAR-2004; 2004AU-00901569.
XX
XX (UYME) UNIV MELBOURNE.
XX
XX Kent SJ;
XX
XX WPI; 2005-031657/03.
XX
XX Use of at least one set of peptides in the preparation of a medicament
XX for modulating an immune response, and for treating cancer or yeast,
XX viral, bacterial, protozoal and mycoplasma infections.
XX
XX Disclosure; SEQ ID NO 1420; 645pp; English.
XX
XX The invention relates to the use of at least one set of peptides in the
XX preparation of a medicament for modulating an immune response, where
XX individual peptides of a respective set comprise different portions of an
XX amino acid sequence corresponding to a single polypeptide of interest and
XX display partial sequence identity or similarity to at least one other
XX peptide of the same set of peptides (i.e. they are overlapping). Also
XX included are an antigen-presenting cell which has been contacted with the
XX peptides above and thus presents the peptides, a population of such
XX antigen-presenting cells, a process for producing antigen-presenting
XX cells for modulating an immune response to a polypeptide of interest, a
XX method for producing antigen-specific lymphocytes, a composition
XX comprising at least one set of the peptides (and a carrier and/or
XX diluent), a method for modulating an immune response to a polypeptide of
XX interest comprising administering to a patient in need at least one set
XX of the peptides, a method for treatment and/or prophylaxis of a disease
XX or condition associated with the presence of a polypeptide of interest
XX and a composition of matter for modulating an immune response in a
XX subject to a target antigen. The polypeptide of interest is also a
XX disease- or condition-associated polypeptide that is a polypeptide
XX produced by a pathogenic organism or a cancer, and produced by a
XX pathogenic organism selected from yeast, viruses, bacteria, helminths,
XX protozoans and mycoplasmas. The disease- or condition-associated
XX polypeptide is produced by a cancer selected from melanoma, lung cancer,
XX breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
XX cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
XX lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured

CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC or aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC C virus protein.

XX SQ Sequence 18 AA;

Query Match 11.0%; Score 13; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 8.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 5 GGVLAALAAAYCLS 17
|||||

RESULT 39

AAR63317
ID AAR63317 standard; protein; 20 AA.

XX AC AAR63317;

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 09-AUG-1995 (first entry)

XX Peptide fragment of hepatitis C virus type 3.

DE Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
XX classification; immunisation; prophylaxis; serotyping.

XX Hepatitis C virus; isolate HD10.

XX WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.

XX Disclosure; Page 166; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
CC genomic sequence, a type 4 genomic sequence; or the coding region of
CC subtype 5a, may be used as primers to amplify nucleic acid from an
CC isolate belonging to a specific genotype, or as a probe for specific
CC detection/classification of nucleic acid. Polypeptides encoded by the
CC nucleotides in such compositions may be used for immunisation against

CC HCV, for the detection of antibodies directed against HCV and for
CC serotyping. This polypeptide corresponds to positions 1688-1707 of HCV
CC type 3. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-
CC 2003 to standardise OS field)

XX SQ Sequence 20 AA;

Query Match 11.0%; Score 13; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQOYDE 61
Db 8 VPDKEVLYQOYDE 20
|||||

RESULT 40

AAR63319
ID AAR63319 standard; protein; 20 AA.

XX AC AAR63319;

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 09-AUG-1995 (first entry)

XX Peptide fragment of hepatitis C virus type 3.

DE Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
XX classification; immunisation; prophylaxis; serotyping.

XX Hepatitis C virus; isolate BR36.

XX WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.

XX Disclosure; Page 167; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
CC genomic sequence, a type 4 genomic sequence; or the coding region of
CC subtype 5a, may be used as primers to amplify nucleic acid from an
CC isolate belonging to a specific genotype, or as a probe for specific
CC detection/classification of nucleic acid. Polypeptides encoded by the
CC nucleotides in such compositions may be used for immunisation against
CC HCV, for the detection of antibodies directed against HCV and for
CC serotyping. This polypeptide corresponds to positions 1724-1743 of HCV
CC type 3. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-
CC 2003 to standardise OS field)

XX SQ Sequence 20 AA;

XX Sequence 24 AA;
SQ

Query Match 11.0%; Score 13; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 9.1e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLLQRATQQQ 97
| | | | | | | |
Db 8 KVLGLLQRATQQQ 20

RESULT 41
AAR41166
ID AAR41166 standard; peptide; 24 AA.
AC AAR41166;
XX
DT 25-MAR-2003 (revised)
DT 22-MAR-1994 (first entry)
XX
DE HCV (type 3) peptide NS4-7 (3).
XX
KW Human immunodeficiency virus; HTV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
XX
OS Synthetic.
XX

FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B=
FT biotin; X= biotinylation cpd. incorporated during
FT synthesis; Y= bond or linking gp(s), which minimises
FT steric hindrance, where Y is not a bond it is pref. 1-10
FT residues of (same or different) glycine, beta-alanine, 4-
FT aminobutyric acid, 5-aminovaleric acid or 6-aminohexanoic
FT acid; parentheses around B and X indicate opt. presence
FT at the specified positions but B or X must be present in
FT at least one of the positions shown, B interacts with the
FT peptide to give a cpd. with greater diagnostic
FT sensitivity; A (optional)= one or more amino acids, NH2
FT or gp. which modifies the N-terminus; Z= one or more
FT amino acids, OH, NH2, or a linkage involving either of
FT these 2 gps."
FT Modified-site 24
FT /note= "the C-terminal comprises Y-(X)-Z"
FT XX
FN W09318054-A2.
XX
PD 16-SEP-1993.
XX
PF 08-MAR-1993; 93WO-EP000517.
XX
PR 06-MAR-1992; 92EP-00400598.
XX
PA (INNO-) INNOGENETICS NV SA.
XX
PI De Leys R;
XX
DR WPI; 1993-303397/38.
XX
PT New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
PT increased antigenicity, useful in antibodies detection and vaccines
PT against hepatitis C, HIV and HTLV.
XX
PS Claim 4; Page 90-98; 133pp; English.

CC Peptide compns. comprise at least one and pref. a combination of two,
CC three, four or more biotinylated peptides chosen from the sequences given
CC in AAR41058-R41166. The peptides represent immunologically important
CC regions of viral proteins and are prepd. by solid phase peptide
CC synthesis. The compns. are useful for the detection of antibodies to
CC HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct
CC CN field.)

XX Query Match 11.0%; Score 13; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLLQRATQQQ 97
| | | | | | | |
Db 11 KVLGLLQRATQQQ 23

RESULT 42
AAU84709
ID AAU84709 standard; peptide; 30 AA.
XX
AC AAU84709;
XX
DT 08-MAY-2002 (first entry)
XX
DE HCV HepC1a segment 112.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Hepatitis C virus.
XX
PN W0200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
FI Thomson SA, Ramshaw IA;
XX WPI; 2002-147575/19.
XX DR N-PSDB; ABK36547.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 2; Fig 26; 364pp; English.

CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcus, Haemophilus,
CC Salmonella, Streptococcus, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a peptide derived from a parent protein used to construct a savine of the

```
CC invention
XX
SQ Sequence 30 AA;

Query Match      11.0%; Score 13; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1 GGVLAALAAAYCLS 13

RESULT 43
AAU84708
ID AAU84708 standard; peptide; 30 AA.
XX
AC AAU84708;
XX
DT 08-MAY-2002 (first entry)
XX
DE HCV HepC1a segment 111.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Hepatitis C virus.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
WPI; 2002-147575/19.
XX
N-PSDB; ABK36546.

New synthetic polypeptides having several different segments of at least
one parent polypeptide linked together differently compared to the
linkage in the parent polypeptide, for inducing immune response against a
pathogen or cancer.

Example 2; Fig 26; 364pp; English.

The invention relates to a new synthetic polypeptide (I) comprising
several different segments of at least one parent polypeptide linked
together in a different relationship relative to their linkage in the
parent polypeptide to impede, abrogate or otherwise alter at least one
function associated with the parent polypeptide and for inducing an
immune response against a pathogen or cancer. Also included are a
synthetic polynucleotide encoding and a computer system for designing the
synthetic polypeptides. The synthetic polypeptides and polynucleotides
are referred to as a Savine. The synthetic polypeptide is useful for
modulating immune responses preferably directed against a pathogen or a
cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
oesophagus, brain, testicle, uterus), as potentiating agents.
Compositions comprising the polypeptide may be used in the treatment or
prophylaxis against viral (such as infections caused by HIV (human
immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
(e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
(e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
```

```
CC a peptide derived from a parent protein used to construct a savine of the
CC invention
XX
SQ Sequence 30 AA;

Query Match      11.0%; Score 13; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 16 GGVLAALAAAYCLS 28

RESULT 44
AAW09248
ID AAW09248 standard; protein; 51 AA.
XX
AC AAW09248;
XX
DT 29-MAR-1997 (first entry)
XX
DE HCV truncated NS4A cofactor (aa 4-54).
XX
KW HCV; NS4A; NS3 protease; substrate; nonstructural polyprotein; inhibitor;
KW assay; liver disease; hepatocellular carcinoma; tumour.
XX
OS Hepatitis C virus.
XX
PN WO9635717-A2.
XX
PD 14-NOV-1996.
XX
PF 09-MAY-1996; 96WO-US006389.
XX
PR 12-MAY-1995; 95US-00439747.
XX
PA (SCHE) SCHERING CORP.
XX
PI Zhang R, Murray MG, Ramanathan L;
XX
WPI; 1996-518617/51.
XX
N-PSDB; AAT42396.

New soluble substrates for hepatitis C virus NS3 protease - are non-
structural polyproteins and are attached to solubilising motifs, useful
for determining protease inhibitors.

Disclosure; Page 61-62; 70pp; English.

Hepatitis C virus (HCV) truncated cofactor NS4A (AAW09248), comprising
amino acids 4-54 of the native cofactor (AAW09292), is used in novel
fusions with HCV NS3 protease (see also AAW09236-40). NS4A improves the
ability of NS3 protease (see also AAW12963) to cleave HCV nonstructural
proteins. The protease-NS4A fusions can be used with substrate peptides
(AAW12957-62) in novel high throughput assays to identify HCV protease
inhibitors of potential therapeutic value

SQ Sequence 51 AA;

Query Match      11.0%; Score 13; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 0.0002;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 4 GGVLAALAAAYCLS 16

RESULT 45
ABB77267
ID ABB77267 standard; protein; 51 AA.
XX
```


AC ABB77267;
 XX 28-JUN-2002 (first entry)
 XX HCV bait polypeptide 15.
 XX
 XX SID; selected interacting domain; HCV; hepatitis C virus; liver disease;
 KW liver cancer; virucide; hepatotropic; antiinflammatory; antibacterial.
 XX
 XX Hepatitis C virus strain H77.
 XX
 XX EP1178116-A1.
 XX
 XX 06-FEB-2002.
 XX
 XX 03-AUG-2000; 2000EP-00402225.
 XX
 XX 03-AUG-2000; 2000EP-00402225.
 XX (HYBR-) HYBRIGENICS SA.
 XX
 XX Legrain P, Whiteside S, Wojcik J;
 XX
 XX WPI; 2002-208115/27.
 DR N-PSDB; ABL55599.
 XX
 XX New selected interacting domain polypeptides and polynucleotides, useful
 PT for treating or preventing infections or pathologies caused by hepatitis
 PT C virus (HCV) or those linked to HCV infection.
 XX
 XX Claim 26; SEQ ID NO 91; 61pp + Sequence Listing; English.
 XX
 XX The invention relates to nucleic acids encoding polypeptides which are
 CC termed SID polypeptides (selected interacting domain). These polypeptides
 CC are the final products of a double selection method involving a first
 CC step of selection of Hepatitis C virus (HCV)-derived polynucleotides
 CC through a two-hybrid system, and a second selection step involving an
 CC alignment between the different polynucleotides selected at the first
 CC step. The activity of polypeptides of the invention may be described as,
 CC virucide, hepatotropic, antiinflammatory and antibacterial. The
 CC polypeptide, polynucleotide and compositions comprising them are useful
 CC for treating or preventing viral or a bacterial infection, specifically
 CC infections or pathologies caused by HCV, or those pathologies linked to
 CC HCV infection. These may include liver disease and liver cancer. The
 CC current sequence represents a HCV bait polypeptide. Note: The sequence
 CC data for this patent is not represented in the specification, but is
 CC based on sequence information supplied by the European Patent Office
 XX
 XX Sequence 51 AA;
 SQ
 Query Match 11.0%; Score 13; DB 5; Length 51;
 Best Local Similarity 100.0%; Pred. No. 0.0002;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 DB 8 GGVLAALAAAYCLS 20
 RESULT 46
 AAW09242
 ID AAW09242 standard; protein; 54 AA.
 XX
 XX AAW09242;
 AC
 XX 29-MAR-1997 (first entry)
 XX
 XX HCV NS4A cofactor.
 DT
 XX
 DE HCV NS4A cofactor.
 XX
 XX HCV; NS4A; NS3 protease; substrate; nonstructural polypeptide; inhibitor;
 KW assay; liver disease; hepatocellular carcinoma; tumour.
 XX
 XX Hepatitis C virus.
 OS

XX WO9635717-A2.
 PN
 XX 14-NOV-1996.
 PD
 XX 09-MAY-1996; 96WO-US006389.
 PF
 XX 12-MAY-1995; 95US-00439747.
 PR
 XX (SCHE) SCHERING CORP.
 PA
 XX Zhang R, Murray MG, Ramanathan L;
 PI WPI; 1996-518617/51.
 XX N-PSDB; AAT42395, AAT42397.
 DR
 XX New soluble substrates for hepatitis C virus NS3 protease - are non-
 PT structural poly:proteins and are attached to solubilising motifs, useful
 PT for determining protease inhibitors.
 XX
 XX Disclosure; Page 45; 70pp; English.
 PS
 XX The hepatitis C virus (HCV) cofactor NS4A (AAW09242) improves the ability
 CC of HCV NS3 protease (see also AAW12963) to cleave HCV nonstructural
 CC proteins. Novel NS3 protease-NS4A fusions (see also AAW09236-40) have
 CC been produced that can be used with NS3 protease substrates (AAW12957-62)
 CC in novel high throughput assays to identify HCV protease inhibitors of
 CC potential therapeutic appln
 XX
 XX Sequence 54 AA;
 SQ
 Query Match 11.0%; Score 13; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 0.00021;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 DB 7 GGVLAALAAAYCLS 19
 RESULT 47
 AAW04579
 ID AAW04579 standard; protein; 54 AA.
 XX
 XX AAW04579;
 AC
 XX 08-FEB-1997 (first entry)
 DT
 XX
 DE Hepatitis C virus NS4A protease.
 XX
 XX Serine protease; NS3; NS4A; HCV; inclusion body; solubilisation;
 KW refolding; renaturation.
 XX
 XX Hepatitis C virus.
 OS
 XX WO9635709-A1.
 PN
 XX 14-NOV-1996.
 PD
 XX 09-MAY-1996; 96WO-US006388.
 PF
 XX 12-MAY-1995; 95US-00439680.
 PR
 XX 13-DEC-1995; 95US-00571643.
 XX
 XX (SCHE) SCHERING CORP.
 PA
 XX Ramanathan L, Wendel M;
 PI WPI; 1996-518613/51.
 XX N-PSDB; AAT43707.
 DR
 XX Prodn. of soluble, active HCV NS3 protease - from insoluble aggregates
 PT produced by bacteria, using denaturing and reducing agent.

XX PS Disclosure; Page 27; 36pp; English.

XX CC The NS4A protease (AAW04579) of hepatitis C virus (HCV) is involved in the cleavage of the non-structural proteins necessary for HCV replication. NS4A or its C-terminal polypeptide (see also AAW04581) can be used to enhance the activity of HCV NS3 protease (see also AAW04578 and AAW04582) following the solubilisation and refolding of NS3 from bacterial inclusion bodies

XX CC bacterial inclusion bodies

XX SQ Sequence 54 AA;

Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 7 GGVLAALAAAYCLS 19

RESULT 48

AAW01651

ID AAW01651 standard; protein; 54 AA.

AC AAW01651;

XX 21-APR-1997 (first entry)

XX HCV NS4A cofactor.

XX HCV; NS4A cofactor; NS3 protease.

XX Hepatitis C virus.

XX W09636702-A2.

XX 21-NOV-1996.

XX 09-MAY-1996; 96WO-US006387.

XX 12-MAY-1995; 95US-00440409.

XX (SCHE) SCHERING CORP.

XX Dasmahapatra B, Murray MG, Ramanathan L, Butkiewicz NJ;

XX WPI; 1997-012081/01.

XX N-PSDB; AAT58402.

XX Bacterially produced Hepatitis C virus NS3 protease(s) - which are denatured and re-folded to produce soluble, active enzyme.

XX Disclosure; Page 46; 71pp; English.

XX The NS4A cofactor protein (AAW01651) of hepatitis C virus (HCV) is a component of novel fusion proteins (see also AAW01646-49) together with the catalytic domain (see also AAW01641) of the HCV NS3 protease. The processing activity of NS3 protease catalytic domain is enhanced by the presence of NS4A, and the activity of fusion proteins contrg. the NS3 protease catalytic domain and NS4A cofactor is superior to that of proteins comprising the NS3 catalytic domain alone

XX SQ Sequence 54 AA;

Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 7 GGVLAALAAAYCLS 19

RESULT 49

AAW47144

ID AAW47144 standard; peptide; 54 AA.

XX AAW47144;

XX 26-MAY-1998 (first entry)

XX HCV NS4A polypeptide fragment.

XX Hepatitis C virus; HCV; HCV NS3 protease; inhibitor; bivalent; monovalent; HCV NS4A; HCV infection.

XX Hepatitis C virus.

XX W09743310-A1.

XX 20-NOV-1997.

XX 08-MAY-1997; 97WO-US007632.

XX 10-MAY-1996; 96US-00644544.

XX (SCHE) SCHERING CORP.

XX Zhang R, Mui PW, Weber PC;

XX WPI; 1998-008797/01.

XX N-PSDB; AAV15929.

XX New inhibitors of hepatitis C protease NS3 - contain at least one of NS3 substrate sequence and NS4A co:factor polypeptide, for treatment of hepatitis C infection.

XX Disclosure; Page 37; 59pp; English.

XX This is a fragment of hepatitis C virus (HCV) NS4A polypeptide. This can be used for assaying the inhibitory activity of novel bivalent inhibitors of HCV NS3 protease. The bivalent inhibitors comprise a first peptide that is a subsequence, mutated subsequence or a mutated full-length sequence of the NS3 substrate linked to a second peptide that is a subsequence of the HCV NS4A polypeptide. These bivalent inhibitors and other monovalent inhibitors of an HCV protease comprising a subsequence, mutated subsequence or a mutated full-length sequence of the substrate of HCV NS3 protease or a subsequence, mutated subsequence or mutated full-length sequence of NS4A are used to treat HCV infection. They act by inhibiting the interaction between NS3 and at least one of its substrates and the NS4A co-factor. Compared with inhibitors that target only one component, the bivalent inhibitors may have higher binding affinity and better discrimination against similar host cell enzymes, i.e. reduced toxicity. The peptide inhibitors can be assessed for their inhibitory activity by a scintillation proximity assay using NS3, NS4 and peptide substrates 4B/5A or 5A/5B. The inhibitors are made by usual methods of solid phase synthesis and can be administered orally or by injection or by transdermal diffusion, optionally conjugated to a carrier protein

XX SQ Sequence 54 AA;

Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 7 GGVLAALAAAYCLS 19

RESULT 50

AAW57203

ID AAW57203 standard; protein; 54 AA.

XX AAW57203;

XX 29-FEB-2000 (first entry)
XX HCV NS4A protein.
XX Hepatitis C virus; HCV; NS3 protease; bivalent inhibitor; linker;
KW NS4A polypeptide; monovalent inhibitor.
XX Hepatitis C virus.
OS US5990276-A.
PN 23-NOV-1999.
XX 09-MAY-1997; 97US-00853623.
XX 10-MAY-1996; 96US-0017470P.
XX (SCHE) SCHERING CORP.
XX Zhang R, Mui PW, Weber PC;
PI WPI; 2000-037868/03.
XX N-PSDB; AA238184.
XX Bivalent and monovalent inhibitors of hepatitis C virus NS3 protease.
PS Disclosure; Col 27-28; 27pp; English.
XX The invention provides bivalent inhibitors of hepatitis C virus (HCV) NS3
CC protease. The bivalent inhibitor comprises: (a) a first peptide
CC consisting of a subsequence, a mutated subsequence or a mutated full-
CC length sequence of a substrate of the HCV NS3 protease which is not
CC cleaved by the protease; (b) a second peptide consisting of a subsequence
CC of a HCV NS4A polypeptide (sequences AAY57195-201); (c) a linker
CC comprising a chemical entity capable of forming a bond with the first
CC peptide and the second peptide and is equivalent in length to a carbon
CC chain having 7-14 carbon atoms. Monovalent inhibitors of the HCV NS3
CC protease inhibit either the interaction of a substrate or the cofactor
CC NS4A with the NS3 protease, and the bivalent inhibitor inhibits the
CC interaction of the NS3 protease with both cofactor NS4A and a substrate
CC of the NS3 protease. The mono- and bivalent inhibitors are useful for
CC treating an individual infected with the HCV. The bivalent enzyme
CC inhibitors provide a higher binding affinity (potency), as well as
CC enhanced specificity against similar cellular host enzymes for reduced
CC toxicity effects. The present sequence represents the HCV NS4A
XX protein
SQ Sequence 54 AA;
Query Match 11.0%; Score 13; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 7 GGVLAALAAAYCLS 19
RESULT 51
AAY57207
ID AAY57207 standard; protein; 54 AA.
XX AAY57207;
XX 29-FEB-2000 (first entry)
XX HCV NS4A native protein.
XX Hepatitis C virus; HCV; NS3 protease; bivalent inhibitor; linker;
KW NS4A polypeptide; monovalent inhibitor.
XX Hepatitis C virus.
OS Hepatitis C virus.

PN US5990276-A.
XX 23-NOV-1999.
XX 09-MAY-1997; 97US-00853623.
XX 10-MAY-1996; 96US-0017470P.
XX (SCHE) SCHERING CORP.
XX Zhang R, Mui PW, Weber PC;
PI WPI; 2000-037868/03.
XX N-PSDB; AA238189.
XX Bivalent and monovalent inhibitors of hepatitis C virus NS3 protease.
PS Disclosure; Col 31-32; 27pp; English.
XX The invention provides bivalent inhibitors of hepatitis C virus (HCV) NS3
CC protease. The bivalent inhibitor comprises: (a) a first peptide
CC consisting of a subsequence, a mutated subsequence or a mutated full-
CC length sequence of a substrate of the HCV NS3 protease which is not
CC cleaved by the protease; (b) a second peptide consisting of a subsequence
CC of a HCV NS4A polypeptide (sequences AAY57195-201); (c) a linker
CC comprising a chemical entity capable of forming a bond with the first
CC peptide and the second peptide and is equivalent in length to a carbon
CC chain having 7-14 carbon atoms. Monovalent inhibitors of the HCV NS3
CC protease inhibit either the interaction of a substrate or the cofactor
CC NS4A with the NS3 protease, and the bivalent inhibitor inhibits the
CC interaction of the NS3 protease with both cofactor NS4A and a substrate
CC of the NS3 protease. The mono- and bivalent inhibitors are useful for
CC treating an individual infected with the HCV. The bivalent enzyme
CC inhibitors provide a higher binding affinity (potency), as well as
CC enhanced specificity against similar cellular host enzymes for reduced
CC toxicity effects. The present sequence represents the native HCV NS4A
XX protein
SQ Sequence 54 AA;
Query Match 11.0%; Score 13; DB 3; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 7 GGVLAALAAAYCLS 19
RESULT 52
AAE13186
ID AAE13186 standard; protein; 54 AA.
XX AAE13186;
XX 28-JAN-2002 (first entry)
XX Hepatitis C virus non-structural (NS) 4A protein.
XX Hepatitis C virus; HCV; GB virus-B; chimeric virus propagation;
KW HCV replication; infection; non-structural protein; NS4A.
XX Hepatitis C virus.
XX Key Location/Qualifiers
FT Region 22..33
FT /label= Cofactor_region
FT Misc-difference 25
FT /note= "Residue critical for the cofactor activation of
FT the NS3 protease"
FT Misc-difference 29
FT /note= "Residue critical for the cofactor activation of
FT the NS3 protease"

```
XX PN US2001034019-A1.
XX PD 25-OCT-2001.
XX PF 21-DEC-2000; 2000US-00742659.
XX PR 22-DEC-1999; 99US-0171469P.
XX PA (HONG/) HONG Z.
XX PA (BUTK/) BUTKIEWICZ N J.
XX PA (ZHON/) ZHONG W.
XX PA (INGR/) INGRAVALLO P.
XX PA (WRIG/) WRIGHT-MINOUE J.
XX PA (LAUJ/) LAU J Y.
XX PA (LEMO/) LEMON S M.
XX PI Hong Z; Butkiewicz NJ, Zhong W, Ingravallo P, Wright-Minogue J;
XX PI Lau JY, Lemon SM;
XX DR WPI; 2001-662529/76.
XX PT New nucleic acids encoding hepatitis C virus (HCV)/GB virus-B (GBV-B)
XX PT constructs for studying HCV replication and infection and for screening
XX PT for HCV inhibitors.
XX PS Example 1; Fig 6; 30pp; English.
XX CC The patent discloses nucleic acids encoding Hepatitis C virus (HCV)/GB
XX CC virus-B (GBV-B) constructs and chimeric, infectious HCV/GBV-B viruses
XX CC comprising non-structural (NS) protein functions of HCV. The HCV/GBV-B
XX CC nucleic acid constructs are used for propagating chimeric viruses. The
XX CC nucleic acid constructs and chimeric viruses are useful for studying HCV
XX CC replication and infection and for screening HCV inhibitors. The present
XX CC sequence is Hepatitis C virus non-structural (NS) 4A protein
XX SQ Sequence 54 AA;
XX Query Match 11.0%; Score 13; DB 4; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 0.00021;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCLS 30
Db |||||
7 GGVLAAALAAAYCLS 19
RESULT 54
AAE21848
ID AAE21848 standard; protein; 54 AA.
XX AC AAE21848;
XX DT 16-JUL-2002 (first entry)
XX DE Hepatitis C virus NS3/4A truncated mutant protein #2.
XX KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX KW virucide; mutant; muten.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FN WO200214362-A2.
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M;
XX PI WPI; 2002-339446/37.
XX DR Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX PT absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Example 1; Page 84; 90pp; English.
```

```
XX PN US2001034019-A1.
XX PD 25-OCT-2001.
XX PF 21-DEC-2000; 2000US-00742659.
XX PR 22-DEC-1999; 99US-0171469P.
XX PA (HONG/) HONG Z.
XX PA (BUTK/) BUTKIEWICZ N J.
XX PA (ZHON/) ZHONG W.
XX PA (INGR/) INGRAVALLO P.
XX PA (WRIG/) WRIGHT-MINOUE J.
XX PA (LAUJ/) LAU J Y.
XX PA (LEMO/) LEMON S M.
XX PI Hong Z; Butkiewicz NJ, Zhong W, Ingravallo P, Wright-Minogue J;
XX PI Lau JY, Lemon SM;
XX DR WPI; 2001-662529/76.
XX PT New nucleic acids encoding hepatitis C virus (HCV)/GB virus-B (GBV-B)
XX PT constructs for studying HCV replication and infection and for screening
XX PT for HCV inhibitors.
XX PS Example 1; Fig 6; 30pp; English.
XX CC The patent discloses nucleic acids encoding Hepatitis C virus (HCV)/GB
XX CC virus-B (GBV-B) constructs and chimeric, infectious HCV/GBV-B viruses
XX CC comprising non-structural (NS) protein functions of HCV. The HCV/GBV-B
XX CC nucleic acid constructs are used for propagating chimeric viruses. The
XX CC nucleic acid constructs and chimeric viruses are useful for studying HCV
XX CC replication and infection and for screening HCV inhibitors. The present
XX CC sequence is Hepatitis C virus non-structural (NS) 4A protein
XX SQ Sequence 54 AA;
XX Query Match 11.0%; Score 13; DB 4; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 0.00021;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCLS 30
Db |||||
7 GGVLAAALAAAYCLS 19
RESULT 53
AAE10069
ID AAE10069 standard; protein; 54 AA.
XX AC AAE10069;
XX DT 29-NOV-2001 (first entry)
XX DE Hepatitis C virus (HCV) NS4a cofactor protein.
XX KW Hepatitis C virus; HCV protease; chromogenic substrate; fluorophore;
XX KW chromophore; fluorogenic; fluorescence polarisation substrate.
XX OS Hepatitis C virus.
XX FN US6251583-B1.
XX PD 26-JUN-2001.
XX PF 08-APR-1999; 99US-00288391.
XX PR 27-APR-1998; 98US-0083204P.
XX PA (SCHE ) SCHERING CORP.
XX PI Zhang R, Malcolm BA, Beyer BM, Njoroge FG, Durkin JP, Windsor WT;
```

Db	7	GGVLAALAAAYCLS 19	
RESULT 56			
AAE19906			
ID	AAE19906	standard; peptide; 54 AA.	
XX	AC		
XX	AAE19906;		
18-JUN-2002	(first entry)		
XX			
XX	Hepatitis C virus (HCV) NS4A peptide.		
DE			
XX	Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;		
KW	cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.		
XX			
OS	Hepatitis C virus.		
XX			
XX	WO200213855-A2.		
PN			
XX			
PD	21-FEB-2002.		
XX			
XX	15-AUG-2001; 2001WO-IB001808.		
PF			
XX			
XX	17-AUG-2000; 2000US-0225767P.		
PR			
PR	29-AUG-2000; 2000US-0229175P.		
PR	03-NOV-2000; 2000US-00705547.		
XX			
XX	(TRIP-) TRIPEP AB.		
PA			
XX			
XX	Sallberg M, Hultgren C;		
PI			
XX	WPI; 2002-241837/29.		
DR			
XX			
PT	Vaccine compositions for treating and preventing disease, preferably		
PT	hepatitis C virus infection, comprises ribavirin and antigen that has		
PT	epitope present in hepatitis C virus.		
XX			
XX	Example 13; Page 101; 120pp; English.		
PS			
XX			
CC	The invention relates to a composition comprising ribavirin and an		
CC	antigen preferably non structural 3 protein (NS3)/4A fragment of		
CC	hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV		
CC	sequence. The composition is useful for enhancing an immune response to a		
CC	hepatitis C antigen in humans, domestic, sport or pet species and as		
CC	vaccines for treating and preventing HCV infections. The composition is		
CC	also useful for treating viral, bacterial, fungal diseases and cancer.		
CC	The present sequence is HCV NS4A peptide		
XX			
SQ	Sequence 54 AA;		
Query Match	11.0%; Score 13; DB 5; Length 54;		
Best Local Similarity	100.0%; Pred. No. 0.00021;		
Matches 13; Conservative	0; Mismatches 0; Indels 0; Gaps 0		
QY	18	GGVLAALAAAYCLS 30	
Db	7	GGVLAALAAAYCLS 19	
RESULT 57			
ABW00345			
ID	ABW00345	standard; peptide; 54 AA.	
XX	AC		
XX	ABW00345;		
AC			
DT	15-JAN-2004 (first entry)		
XX			
XX	Hepatitis C virus NS4A peptide.		
DE			
XX	Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;		
KW	virucide.		
XX			

```
OS Hepatitis C virus.
XX
XX US2002136740-A1.
XX
XX 26-SEP-2002.
XX
XX 15-AUG-2001; 2001US-00929955.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Claim 11; Page 40; Opp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS4A peptide
XX
XX Sequence 54 AA;
XX
XX Query Match 11.0%; Score 13; DB 7; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 0.00021;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX Db 7 GGVLAALAAAYCLS 19
XX
XX RESULT 59
XX ADG47670
XX ID ADG47670 standard; protein; 54 AA.
XX
XX AC ADG47670;
XX
XX DT 11-MAR-2004 (first entry)
XX
XX DE HCV NS3/4A domain mutant #11.
XX
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX
XX OS Synthetic.
XX OS Hepatitis C virus.
XX
XX PN US2003206919-A1.
XX
XX PD 06-NOV-2003.
XX
XX PF 26-NOV-2002; 2002US-00307047.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX
XX PA (SALL/) SALLBERG M.
XX
XX PI Sallberg M;
XX
XX DR WPI; 2004-051480/05.
XX
XX PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX
XX PS Example 1; SEQ ID NO 13; 83pp; English.
XX
XX The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX
XX Sequence 54 AA;
XX
```

```
OS Hepatitis C virus.
XX
XX US2002136740-A1.
XX
XX 26-SEP-2002.
XX
XX 15-AUG-2001; 2001US-00929955.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Claim 11; Page 40; Opp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS4A peptide
XX
XX Sequence 54 AA;
XX
XX Query Match 11.0%; Score 13; DB 7; Length 54;
XX Best Local Similarity 100.0%; Pred. No. 0.00021;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX Db 7 GGVLAALAAAYCLS 19
XX
XX RESULT 58
XX ABW00357
XX ID ABW00357 standard; peptide; 54 AA.
XX
XX AC ABW00357;
XX
XX DT 15-JAN-2004 (first entry)
XX
XX DE Hepatitis C virus NS4A mutant peptide.
XX
XX KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
KW virucide; mutant; mutein.
XX
XX OS Hepatitis C virus.
XX OS Synthetic.
XX
XX PN US2002136740-A1.
XX
XX PD 26-SEP-2002.
XX
XX PF 15-AUG-2001; 2001US-00929955.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX
```

```
Query Match      11.0%; Score 13; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 7 GGVLAALAAAYCLS 19

RESULT 60
AAP92020
ID AAP92020 standard; protein; 135 AA.
XX AC
XX DT 25-MAR-2003 (revised)
XX DT 01-NOV-1989 (first entry)
XX DE Sequence of hepatitis cDNA insert in DNA 36.
XX DE Hepatitis C virus; DNA 36; probe; vaccine.
XX KW Pan troglodytes.
XX OS
XX PN GB2212511-A.
XX PD 26-JUL-1989.
XX PF 18-NOV-1988; 88GB-00027024.
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 26-OCT-1988; 88US-00263584.
XX PA (CHIR ) CHIRON CORP.
XX PI Houghton M, Choo QL, Kuo G;
XX DR WPI; 1989-215054/30.
XX DR N-PSDB; AAN90306.
XX PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
XX PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
XX PS infection.
XX PS Disclosure; Fig 5; 30pp; English.
XX CC The sequence is the peptide encoded by the hepatitis C virus (HCV) cDNA
XX CC insert in DNA 36 (see AAN90306). The polypeptides are used to diagnose
XX CC HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or
XX CC to produce vaccines. Residues 121-135 overlap with DNA 36. (Updated on 25
XX CC -MAR-2003 to correct PR field.)
XX SQ Sequence 135 AA;

Query Match      11.0%; Score 13; DB 1; Length 135;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 96 GGVLAALAAAYCLS 108

RESULT 62
AAW09240
ID AAW09240 standard; protein; 234 AA.
XX AC
XX AC AAW09240;
XX DT 29-MAR-1997 (first entry)
XX DE HCV insoluble NS3 protease-NS4A cofactor fusion.
XX KW HCV; NS3 protease; NS4A; substrate; nonstructural polyprotein; inhibitor;
XX KW assay; liver disease; hepatocellular carcinoma; tumour.
XX OS Hepatitis C virus.
XX EH Key Location/Qualifiers
XX FT 1. .182
XX FT /label= NS3_catalytic domain
XX FT /note= "amino acids 1-182 of HCV NS3 protease"
XX FT 184. .234
```

```

FT FT /label= NS4A
FN /note= "amino acids 4-54 of HCV NS4A"
PN
PD WO9635717-A2.
XX 14-NOV-1996.
XX
PF 09-MAY-1996; 96WO-US006389.
XX
PR 12-MAY-1995; 95US-00439747.
XX
PA (SCHE ) SCHERING CORP.
XX
PI Zhang R, Murray MG, Ramanathan L;
XX WPI; 1996-518617/51.
XX DR N-PSDB; AAT42393.
XX
XX New soluble substrates for hepatitis C virus NS3 protease - are non-
FT structural poly:proteins and are attached to solubilising motifs, useful
FT for determining protease inhibitors.
XX
XX Disclosure; Page 46-47; 70pp; English.
XX
CC A novel fusion protein (AAW09239) comprises the catalytic domain (see
CC also AAW12963) of hepatitis C virus (HCV) protease NS3 (NS3p) fused to
CC cofactor NS4A (see also AAW09249). It can be produced in transformed host
CC cells using a vector incorporating the encoding cDNA sequence (AAW42393).
CC Purification is facilitated by incorporation of an N-terminal His tag.
CC The NS3p-NS4A fusion can be used with NS3p peptide substrates (AAW12957-
CC 62) in novel high throughput assays to identify HCV protease inhibitors
CC of potential therapeutic appln
XX
SQ Sequence 234 AA;

Query Match 11.0%; Score 13; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 187 GGVLAALAAAYCLS 199

RESULT 63
AAW01649
ID AAW01649 standard; protein; 234 AA.
XX
AC AAW01649;
XX
DT 27-AUG-2003 (revised)
DT 21-APR-1997 (first entry)
XX
DE HCV NS3 protease catalytic domain-NS4A fusion.
XX
XX HCV; NS3 protease; NS4A cofactor; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT 1.182
FT Domain /label= Catalytic domain
FT /note= "amino acid residues 1-182 correspond to amino
FT acids 1-182 of the catalytic domain of HCV NS3 protease"
FT 184..234
FT Domain /label= NS4A
FT /note= "amino acid residues 184-234 correspond to amino
FT acids 4-54 of the NS4A cofactor; NS4A enhances the
FT processing activity of the NS3 protease catalytic domain"
XX
PN WO9636702-A2.

XX PD 21-NOV-1996.
XX
XX 09-MAY-1996; 96WO-US006387.
XX
XX 12-MAY-1995; 95US-00440409.
XX
XX (SCHE ) SCHERING CORP.
XX
XX Dasmahapatra B, Murray MG, Ramanathan L, Butkiewicz NJ;
XX WPI; 1997-012081/01.
XX DR N-PSDB; AAT58400.
XX
XX Bacterially produced Hepatitis C virus NS3 protease(s) - which are
FT denatured and re-folded to produce soluble, active enzyme.
XX
XX Disclosure; Page 47-48; 71pp; English.
XX
CC A fusion protein (AAW01649) is composed of the catalytic domain (see also
CC AAW01641) of hepatitis C virus (HCV) NS3 protease and the HCV NS4A
CC cofactor (see also AAW01651). A cDNA clone (AAT58400) can be used to
CC express large quantities of the polypeptide in E. coli host cells. The
CC fusion protein is produced as an insoluble aggregate, which can be
CC denatured and refolded to produce soluble, active enzyme. Soluble NS3
CC proteases (see also AAW01643-50) are useful in high throughput screens
CC for the detection of inhibitors of the protease that can be used as
CC therapeutic agents against HCV and also for structural studies. (Updated
CC on 27-AUG-2003 to correct OS field.)
XX
XX SQ Sequence 234 AA;

Query Match 11.0%; Score 13; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 187 GGVLAALAAAYCLS 199

RESULT 64
AAP92021
ID AAP92021 standard; protein; 237 AA.
XX
AC AAP92021;
XX
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Polypeptide encoded in the combined ORFs of clones 36 and 81 of HCV.
XX
XX Hepatitis C virus (HCV); non-A, non-B hepatitis (HANEH).
XX
OS Hepatitis C virus.
OS Unidentified.
XX
XX EP318216-A.
XX
XX 31-MAY-1989.
XX
XX 18-NOV-1988; 88EP-00310922.
XX
XX 18-NOV-1987; 87US-00122714.
XX 30-DEC-1987; 87US-00139886.
XX 26-FEB-1988; 88US-00161072.
XX 06-MAY-1988; 88US-00191263.
XX 26-OCT-1988; 88US-00263584.
XX 14-NOV-1988; 88US-00271450.
XX
XX (CHIR ) CHIRON CORP.
XX (CHIR ) CHIRON CORP.

```


CC A novel fusion protein (AAW09238) c
CC also AAW12963) of hepatitis C virus
CC cofactor NS4A (see also AAW09249) w
CC also AAW09246). It can be produced
CC vector incorporating the encoding c
CC solubilised NS3-NS4A fusion can be

A novel fusion protein (AAW09238) comprises the catalytic domain of hepatitis C virus (HCV) protease NS3 (NS3p) and the NS3 cofactor NS4A (see also AAW09249) with a C-terminal hydrophilic tag (see also AAW09246). It can be produced in transformed host cells using a vector incorporating the encoding cDNA sequence (AAW24391). The solubilised NS3-NS4A fusion can be used with NS3 peptide substrates.

```
CC (AAW12957-62) in novel high throughput assays to identify HCV protease
CC inhibitors of potential therapeutic appin
XX
SQ Sequence 237 AA;

Query Match 11.0%; Score 13; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 187 GGVLAALAAAYCLS 199

RESULT 67
AAW01647
ID AAW01647 standard; protein; 237 AA.
XX
AC AAW01647;
XX
DT 27-AUG-2003 (revised)
DT 21-APR-1997 (first entry)
XX
DE HCV NS3 soluble protease catalytic domain-NS4A fusion.
XX
KW HCV; NS3 protease; NS4A cofactor; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Domain 1..182
FT /label= Catalytic domain
FT /note= "amino acid residues 1-182 correspond to amino
FT acids 1-182 of the catalytic domain of HCV NS3 protease"
FT 184..233
FT Domain
FT /label= NS4A
FT /note= "amino acid residues 184-233 correspond to amino
FT acids 4-53 of the NS4A cofactor; NS4A enhances the
FT processing activity of the NS3 protease catalytic domain"
FT 234..237
FT Region
FT /label= Hydrophilic_tail
FT /note= "the C-terminal hydrophilic tail improves the
FT solubility of the fusion protein"
XX
PN WO9636702-A2.
XX
PD 21-NOV-1996.
XX
PF 09-MAY-1996; 96WO-US006387.
XX
PR 12-MAY-1995; 95US-00440409.
XX
PA (SCHE ) SCHERING CORP.
XX
PI Dasamahatra B, Murray MG, Ramanathan L, Butkiewicz NJ;
PI WPI; 1997-012081/01.
DR N-PSDB; AAT58398.
XX
PT Bacterially produced Hepatitis C virus NS3 protease(s) - which are
PT denatured and re-folded to produce soluble, active enzyme.
XX
PS Disclosure; Page 51-53; 71pp; English.
XX
CC A fusion protein (AAW01647) is composed of the catalytic domain (see also
CC AAW01641) of hepatitis C virus (HCV) NS3 protease and the HCV NS4A
CC cofactor (see also AAW01651), plus a C-terminal hydrophilic tail. A cDNA
CC clone (AAT58398) can be used to express large quantities of the
CC polypeptide in E. coli host cells. Soluble, active NS3 proteases (see
CC also AAW01643-50) are useful in high throughput screens for the detection
CC of inhibitors of the protease that can be used as therapeutic agents
```

```
CC against HCV and also for structural studies. (Updated on 27-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 237 AA;

Query Match 11.0%; Score 13; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.00072;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 187 GGVLAALAAAYCLS 199

RESULT 68
AAW09239
ID AAW09239 standard; protein; 250 AA.
XX
AC AAW09239;
XX
DT 29-MAR-1997 (first entry)
DT
XX
DE HCV solubilised NS3 protease-NS4A cofactor fusion.
XX
KW HCV; NS3 protease; NS4A; substrate; nonstructural polyprotein; inhibitor;
KW assay; liver disease; hepatocellular carcinoma; tumour.
XX
OS Hepatitis C virus; virus.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 1..13
FT /label= His_tag
FT Domain 14..195
FT /label= NS3 catalytic domain
FT /note= "amino acids 2-183 of HCV NS3 protease"
FT 197..246
FT Domain
FT /label= NS4A
FT /note= "amino acids 4-53 of HCV NS4A"
FT 248..250
FT Region
FT /label= Solubilising_motif
XX
PN WO9635717-A2.
XX
PD 14-NOV-1996.
XX
PF 09-MAY-1996; 96WO-US006389.
XX
PR 12-MAY-1995; 95US-00439747.
XX
PA (SCHE ) SCHERING CORP.
XX
PI Zhang R, Murray MG, Ramanathan L;
PI WPI; 1996-518617/51.
DR N-PSDB; AAT42392.
XX
PT New soluble substrates for hepatitis C virus NS3 protease - are non-
PT structural polyproteins and are attached to solubilising motifs, useful
PT for determining protease inhibitors.
XX
PS Disclosure; Page 52-54; 70pp; English.
XX
CC A novel fusion protein (AAW09239) comprises the catalytic domain (see
CC also AAW12963) of hepatitis C virus (HCV) protease NS3 (NS3p) fused to
CC cofactor NS4A (see also AAW09249) with a C-terminal hydrophilic tail (see
CC also AAW09246) and an N-terminal His tag to facilitate purification. It
CC can be produced in transfected host cells using a vector incorporating
CC the encoding cDNA sequence (AAT42392). The solubilised NS3p-NS4A fusion
CC can be used with NS3p peptide substrates (AAW12957-62) in novel high
CC throughput assays to identify HCV protease inhibitors of potential
CC therapeutic appin
```

XX Sequence 250 AA;
SQ Query Match 11.0%; Score 13; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
Db 200 GGVLAALAAAYCLIS 212
|||||

RESULT 69
AAW01648
ID AAW01648 standard; protein; 250 AA.
XX AC AAW01648;
XX DT 27-AUG-2003 (revised)
XX DT 21-APR-1997 (first entry)
XX HCV NS3 soluble protease catalytic domain-NS4A fusion.
XX HCV; NS3 protease; NS4A cofactor; inhibitor.
XX HCV; NS3 protease; NS4A cofactor; inhibitor.
XX Hepatitis C virus.
XX Synthetic.
XX Chimeric.
XX Key Location/Qualifiers
XX Region 1..13
FT /label= Hexahistidine tag
FT /note= "the hexahistidine tag facilitates affinity
FT purification of the fusion protein"
FT 14..195
FT /label= Catalytic domain
FT /note= "amino acid residues 14-195 correspond to amino
FT acids 1-182 of the catalytic domain of HCV NS3 protease"
FT 197..246
FT /label= NS4A
FT /note= "amino acid residues 197-246 correspond to amino
FT acids 4-53 of the NS4A cofactor; NS4A enhances the
FT processing activity of the NS3 protease catalytic domain"
FT 247..250
FT /label= Hydrophilic tail
FT /note= "the C-terminal hydrophilic tail improves the
FT solubility of the fusion protein"
XX WO9636702-A2.
XX 21-NOV-1996.
XX 09-MAY-1996; 96WO-US006387.
XX 12-MAY-1995; 95US-00440409.
XX (SCHE) SCHERING CORP.
XX Daemahapatra B, Murray MG, Ramanathan L, Butkiewicz NJ;
XX WPI; 1997-012081/01.
XX N-PSDB; AAT58399.
XX Bacterially produced Hepatitis C virus NS3 protease(s) - which are
XX denatured and re-folded to produce soluble, active enzyme.
XX Disclosure; Page 53-55; 71pp; English.
XX A fusion protein (AAW01648) is composed of the catalytic domain (see also
XX AAW01641) of hepatitis C virus (HCV) NS3 protease and the HCV NS4A
XX cofactor (see also AAW01651), plus a C-terminal hydrophilic tail and N-
XX terminal hexahistidine affinity tag. A cDNA clone (AAT58399) can be used
XX to express large quantities of the polypeptide in E. coli host cells.

CC Soluble, active NS3 proteases (see also AAW01643-50) are useful in high
CC throughput screens for the detection of inhibitors of the protease that
CC can be used as therapeutic agents against HCV and also for structural
CC studies. (Updated on 27-AUG-2003 to correct OS field.)
XX Sequence 250 AA;
SQ Query Match 11.0%; Score 13; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 0.00075;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
Db 200 GGVLAALAAAYCLIS 212
|||||

RESULT 70
ADT77851
ID ADT77851 standard; protein; 252 AA.
XX AC ADT77851;
XX DT 13-JAN-2005 (first entry)
XX Hepatitis C virus non-structural protein NS4.
XX HCV; non-structural protein; NS4; adjuvant; antiinflammatory;
XX immunosuppressive; antimicrobial; cytostatic; antiasthmatic;
XX antiallergic; neuroprotective; antidiabetic; antirheumatic;
XX antiarthritic; dermatological ; ophthalmological.
XX Hepatitis C virus.
XX WO2004089978-A2.
XX 21-OCT-2004.
XX 08-APR-2004; 2004WO-IE0000054.
XX 11-APR-2003; 2003IE-00000279.
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX Mills KHG, Brady MT;
XX WPI; 2004-748721/73.
XX Therapeutic composition for preventing or treating inflammatory or immune
XX -mediated disorders comprises a Hepatitis C virus (HCV) protein activated
XX by an agent that suppresses inflammatory cytokine or promotes IL-10
XX production.
XX Disclosure; SEQ ID NO 1; 43pp; English.
XX The present sequence is the protein sequence of hepatitis C virus (HCV)
XX genotype 1b non-structural protein NS4 comprising amino acid residues
XX 1616-1862 of the HCV polyprotein. The invention provides a method for the
XX treatment and/or prophylaxis of an inflammatory and/or immune-mediated
XX disorder and/or disorders associated with transplantation comprising the
XX step of administering an HCV protein (NS4 or NS3) or its derivative,
XX mutant, fragment or variant, or the product of cells activated thereby.
XX In one case, the HCV protein or its product suppresses inflammatory
XX cytokine production and also promotes interleukin-10 production.
XX particularly by peripheral blood mononuclear cells and/or monocytes. It
XX may also inhibit dendritic cell activation, inhibit the induction or
XX activation of Th1 or Th2 cells, modulate toll-like receptor ligand-
XX induced NFkappaB activation, or modulate inflammatory cytokine production
XX induced by infection or trauma. A claimed therapeutic composition and a
XX claimed vaccine adjuvant comprise HCV NS4 or its derivative mutant, a
XX fragment or variant or the product of cells activated thereby. It can be
XX used in the treatment or prophylaxis of sepsis or acute inflammation
XX induced by infection, trauma or injury, chronic inflammatory disease,
XX graft rejection or graft versus host disease, an immune mediated disease

CC involving Th1 responses, a disease or condition involving toll-like
CC receptor dependent signaling, an immune mediated disease involving
CC inflammatory cytokines including tumour necrosis factor-alpha and
CC interleukin-1, Crohn's disease, inflammatory bowel disease, multiple
CC sclerosis, type 1 diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, uveitis, allergy or asthma (all claimed).
XX
SQ Sequence 252 AA;
Query Match 11.0%; Score 13; DB 8; Length 252;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 52 GGVLAALAAAYCLS 64
RESULT 71
AAW09237
ID AAW09237 standard; protein; 255 AA.
XX
AC AAW09237;
XX
DT 29-MAR-1997 (first entry)
XX
DE HCV solubilised NS3 protease-NS4A cofactor fusion.
XX
KW HCV; NS3 protease; NS4A; substrate; nonstructural polyprotein; inhibitor;
XX assay; liver disease; hepatocellular carcinoma; tumour.
XX
OS Hepatitis C virus; virus.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 1..13
FT /label= His_tag
FT Domain 14..194
FT /label= NS3 catalytic domain
FT /note= "amino acids 2-182 of HCV NS3 protease"
FT Domain 196..246
FT /label= NS4A
FT /note= "amino acids 4-54 of HCV NS4A"
FT Region 247..255
FT /label= Solubilising_motif
XX
PN WO9635717-A2.
XX
PD 14-NOV-1996.
XX
PP 09-MAY-1996; 96WO-US006389.
XX
PR 12-MAY-1995; 95US-00439747.
XX
PA (SCHE) SCHERING CORP.
XX
PI Zhang R, Murray MG, Ramanathan L;
XX
PI WPI; 1996-518617/51.
XX
DR N-PSDB; AAT42390.
XX
PT New soluble substrates for hepatitis C virus NS3 protease - are non-
PT structural poly:proteins and are attached to solubilising motifs, useful
PT for determining protease inhibitors.
XX
PS Example 1; Page 62-64; 70pp; English.
XX
CC A novel fusion protein (AAW09237) comprises the catalytic domain (see
CC also AAW12963) of hepatitis C virus (HCV) protease NS3 (NS3p) fused to
CC cofactor NS4A (see also AAW09249) with a C-terminal solubilising motif
CC (see also AAW09245) and an N-terminal His tag to facilitate purification.
CC It can be produced in transformed host cells using a vector incorporating

CC the encoding cDNA sequence (AAT42390). The solubilised NS3p-NS4A fusion
CC can be used with NS3p peptide substrates (AAW12957-62) in novel high
CC throughput assays to identify HCV protease inhibitors of potential
CC therapeutic appin
XX
SQ Sequence 255 AA;
Query Match 11.0%; Score 13; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 199 GGVLAALAAAYCLS 211
RESULT 72
AAW04572
ID AAW04572 standard; protein; 255 AA.
XX
AC AAW04572;
XX
DT 27-AUG-2003 (revised)
DT 09-FEB-1997 (first entry)
XX
DE HCV NS3-NS4A protease fusion.
XX
KW Protease; inhibitor; surface plasmon resonance; SPR; assay;
XX hepatitis C virus; HCV; pNB(-V)182-delta4A HT; NS3; NS4A.
XX
OS Hepatitis C virus.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Peptide 5..10
FT /note= "hexahistidine affinity tag"
FT Domain 14..194
FT /label= NS3
FT /note= "NS3 protease catalytic domain"
FT Domain 196..245
FT /label= NS4A
FT /note= "NS4A protease amino acids 4-54"
XX
PN WO9635806-A1.
XX
PD 14-NOV-1996.
XX
PP 09-MAY-1996; 96WO-US006385.
XX
PR 12-MAY-1995; 95US-00440283.
XX
PA (SCHE) SCHERING CORP.
XX
PI Taremi SS, Prorise WW;
XX
PI WPI; 1996-518694/51.
XX
DR N-PSDB; AAT38904.
XX
PT Detecting substrate cleavage by enzyme using surface plasmon resonance -
PT with binding agent specific for substrate immobilised on sensor, also
PT used to identify enzyme, specifically protease, inhibitors.
XX
PS Example 3; Page 52-54; 63pp; English.
XX
CC A recombinant polypeptide (AAW04572) comprises the NS3 protease (amino
CC acids 1-182) of hepatitis C virus (HCV) linked to the NS4A protease
CC (amino acids 5-54). It does not contain an NS3/NS4A cleavage site at the
CC junction and is not cleaved by the autocatalytic activity of NS3. It is
CC encoded by cDNA construct pNB(-V)182-delta4A HT (AAT38904). The fusion
CC protein, produced in E. coli transformants, can be used in novel high-
CC throughput assays for inhibitors of HCV protease. Surface plasmon
CC resonance is used to determine whether cleavage of a peptide substrate

CC (see also AAW04561-62) of the protease is inhibited by a test cpd.
CC (Updated on 27-AUG-2003 to correct OS field.)

SQ Sequence 255 AA;

Query Match	11.0%	Score 13;	DB 2;	Length 255;
Best Local Similarity	100.0%	Pred. No. 0.00077;		
Matches 13;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	18	GGVLAALAAAYCLS	30
Dd	199	GGVLAALAAAYCLS	211

RESULT 73
AAW01646
ID AAW01646 standard; protein; 255 AA.

XX	
AC	
XX	AAW01646;
DT	27-AUG-2003 (revised)
DT	21-APR-1997 (first entry)

XX	HCV NS3 soluble protease-NS4A fusion.
DE	
XX	
KW	HCV; NS3 protease; NS4A cofactor; inhibitor.
XX	
OS	Hepatitis C virus.
OS	Synthetic.
OS	Chimeric.

FH	Key	Location/Qualifiers	
FT	Region	1. .13	
FT		/label= Hexahistidine tag	
FT		/note= "the hexahistidine tag facilitates affinity	

```
FT Domain
14..194
/label= Catalytic domain
```

FT	Domain	acids 2-182 of the catalytic domain of HCV NS3 protease"
FT	196..246	
FT	/label = NS4A	
PT		

F1 PT FT
/note= "amino acid residues 195-246 correspond to amino acids 4-54 of the NS4A cofactor; NS4A enhances the processing activity of the NS3 protease catalytic domain"

CC NS4A polypeptide (see also AA001051) plus a C-terminal solubilising motif
CC (see also AA001042) and N-terminal hexahistidine tag. A cDNA clone
CC (AA059397) can be used to express large quantities of the polypeptide in
CC E. coli host cells. Soluble, active NS3 proteases (see also AA001643-50)
CC are useful in high throughput screens for the detection of inhibitors of
CC the protease that can be used as therapeutic agents against HCV and also
CC for structural studies. (Updated on 27-AUG-2003 to correct OS field.)
CC

Sequence 255 AA;

Query Match 11.0%; Score 13; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 0.00077;
Matches 13; Conservative 0; Mismatches 0; Indels

Qy 18 GGVLAAALAAAYCLS 30
|||
Db 199 GGVLAAALAAAYCLS 211

RESULT 74
AAW47147
ID AAW47147 standard; protein; 255 AA.

AC	AAW47147;
XX	
DT	26-MAY-1998 (first entry)
XX	
DE	Chloramphenicol acetyl transferase (CAT) partial sequence.

KW Hepatitis C virus; HCV; HCV NS3 protease; inhibitor; bivalent;
KW monovalent; HCV NS4A co-factor; chloramphenicol acetyl transferase
KW HCV infection; CAT.

PT New inhibitors of hepatitis C protease NS3 - contain at least one of NS3
PT substrate sequence and NS4A co:factor polypeptide, for treatment of
PT hepatitis C infection.

PS Example 3; Page 39-41; 59pp: English.

CC This is a partial sequence of the chloramphenicol acetyl transferase
CC (CAT) gene. This is used in the construction of a plasmid which produces
CC a hepatitis C virus (HCV) NS3 protease. Novel bivalent inhibitors of HCV
CC NS3 protease comprise a first peptide that is a subsequence, mutated
CC subsequence or a mutated full-length sequence of the NS3 substrate linked
CC to a second peptide that is a subsequence of the HCV NS4A polypeptide.

These bivalent inhibitors and other monovalent inhibitors of an HCV protease comprising a subsequence, mutated subsequence or a mutated full-length sequence of the substrate of HCV NS3 protease or a subsequence, mutated subsequence or mutated full-length sequence of NS4A are used to treat HCV infection. They act by inhibiting the interaction between NS3 and at least one of its substrates and the NS4A co-factor. Compared with inhibitors that target only one component, the bivalent inhibitors may have higher binding affinity and better discrimination against similar host cell enzymes, i.e. reduced toxicity. The peptide inhibitors can be assessed for their inhibitory activity by a scintillation proximity assay using NS3, NS4 and peptide substrates 4B/5A or 5A/5B. The inhibitors are

CC made by usual methods of solid phase synthesis and can be administered
 CC orally or by injection or by transdermal diffusion, optionally conjugated
 CC to a carrier protein
 XX
 SQ Sequence 255 AA;

Query Match 11.0%; Score 13; DB 2; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.00077;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||
 Db 199 GGVLAALAAAYCLS 211

RESULT 75
 AAY57206
 ID AAY57206 standard; protein; 255 AA.

XX AAY57206;

DT 29-FEB-2000 (first entry)

XX Amino acid sequence of plasmid pNB1824HT.

XX Hepatitis C virus; HCV; NS3 protease; bivalent inhibitor; linker;
 KW NS4A polypeptide; monovalent inhibitor.

XX Synthetic.

XX US5990276-A.

XX 23-NOV-1999.

XX 09-MAY-1997; 97US-00853623.

XX 10-MAY-1996; 96US-0017470P.

XX (SCHE) SCHERING CORP.

XX Zhang R, Mui PW, Weber PC;

DR WPI; 2000-037868/03.

DR N-PSDB; AAZ38188.

XX Bivalent and monovalent inhibitors of hepatitis C virus NS3 protease.

XX Example 3; Col 29-32; 27pp; English.

XX The invention provides bivalent inhibitors of hepatitis C virus (HCV) NS3
 CC protease. The bivalent inhibitor comprises: (a) a first peptide
 CC consisting of a subsequence, a mutated subsequence or a mutated full-
 CC length sequence of a substrate of the HCV NS3 protease which is not
 CC cleaved by the protease; (b) a second peptide consisting of a subsequence
 CC of a HCV NS4A polypeptide (sequences AA57195-201); (c) a linker
 CC comprising a chemical entity capable of forming a bond with the first
 CC peptide and the second peptide and is equivalent in length to a carbon
 CC chain having 7-14 carbon atoms. Monovalent inhibitors of the HCV NS3
 CC protease inhibit either the interaction of a substrate or the cofactor
 CC NS4A with the NS3 protease, and the bivalent inhibitor inhibits the
 CC interaction of the NS3 protease with both cofactor NS4A and a substrate
 CC of the NS3 protease. The mono- and bivalent inhibitors are useful for
 CC treating an individual infected with the HCV. The bivalent enzyme
 CC inhibitors provide a higher binding affinity (potency), as well as
 CC enhanced specificity against similar cellular host enzymes for reduced
 CC toxicity effects

XX Sequence 255 AA;

Query Match 11.0%; Score 13; DB 3; Length 255;
 Best Local Similarity 100.0%; Pred. No. 0.00077;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||
 Db 199 GGVLAALAAAYCLS 211

RESULT 76

AAW71272

ID AAW71272 standard; protein; 258 AA.

XX AAW71272;

DT 17-NOV-1998 (first entry)

XX Amino acid sequence of a NS3/4A fusion protein of the invention.

XX HCV NS3 protease; NS4A cofactor protein; HCV therapy; fusion protein.

XX Synthetic.

OS Hepatitis C virus.

XX WO9837180-A2.

XX 27-AUG-1998.

XX 20-FEB-1998; 98WO-US003367.

XX 22-FEB-1997; 97US-00804266.

XX (ABBO) ABBOTT LAB.

XX Chen C, Molla A, Tripathi RL;

DR WPI; 1998-467551/40.

DR N-PSDB; AAV54956.

XX New hepatitis C virus fusion proteins - comprises NS3 protease and NS4A
 PT co-factor, used in assays for screening for compounds for use in HCV
 PT therapy.

XX Claim 16; Fig 3; 31pp; English.

XX The present sequence represents a fusion protein of the invention. The
 CC fusion protein is derived from the Hepatitis C virus (HCV) NS3 protease
 CC and NS4A cofactor proteins. A non-autocleavable fusion protein of HCV NS3
 CC protease and HCV NS4A cofactor protein is produced upon expression, which
 CC is biologically active. The products can be used to obtain drugs which
 CC can inhibit NS3 protease activity for use in HCV therapy. They can also
 CC be used to design compounds which interact with or inhibit the NS3/NS4A
 CC fusion proteins

XX Sequence 258 AA;

Query Match 11.0%; Score 13; DB 2; Length 258;
 Best Local Similarity 100.0%; Pred. No. 0.00077;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||

Db 211 GGVLAALAAAYCLS 223

RESULT 77

AAW09236

ID AAW09236 standard; protein; 270 AA.

XX AAW09236;

XX 29-MAR-1997 (first entry)

XX HCV solubilised NS3 protease-NS4A cofactor fusion.

XX HCV, NS3 protease; NS4A; substrate; nonstructural polyprotein; inhibitor;
 KW assay; liver disease; hepatocellular carcinoma; tumour.

```

XX OS Hepatitis C virus; virus.
OS Synthetic.
OS Chimeric.
XX FH Key
XX FT Region
XX FT Location/Qualifiers
XX FT 1..27
XX FT /label= His_tag
XX FT 28..209
XX FT /label= NS3 catalytic domain
XX FT /note= "amino acids 1-182 of HCV NS3 protease"
XX FT 211..261
XX FT /label= NS4A
XX FT /label= NS4A
XX FT /note= "amino acids 4-54 of HCV NS4A"
XX FT 261..270
XX FT /label= Solubilising_motif
XX PN WO9635717-A2.
XX PD 14-NOV-1996.
XX PF 09-MAY-1996; 96WO-US006389.
XX PR 12-MAY-1995; 95US-00439747.
XX PA (SCHE ) SCHERING CORP.
XX PI Zhang R, Murray MG, Ramanathan L;
XX DR WPI; 1996-518617/51.
XX DR N-PSDB; AAT42389.
XX PT New soluble substrates for hepatitis C virus NS3 protease - are non-
XX PT structural poly:proteins and are attached to solubilising motifs, useful
XX PT for determining protease inhibitors.
XX PS Example 1; Page 48-50; 70pp; English.
XX CC A novel fusion protein (AAW09236) comprises the catalytic domain (see
XX CC also AAW12963) of hepatitis C virus (HCV) protease NS3 fused (NS3p) to
XX CC cofactor NS4A (see also AAW09249) with a C-terminal solubilising motif
XX CC (see also AAW09245) and an N-terminal His tag to facilitate purification.
XX CC It can be produced in transformed host cells using a vector incorporating
XX CC the encoding cDNA sequence (AAT42389). The solubilised NS3p-NS4A fusion
XX CC can be used with NS3p substrates (AAW12957-62) in novel high throughput
XX CC assays to identify HCV protease inhibitors of potential therapeutic appln
XX SQ Sequence 270 AA;
XX Query Match 11.0%; Score 13; DB 2; Length 270;
XX Best Local Similarity 100.0%; Pred. No. 0.0008;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAALAAYCLS 30
XX DB 214 GGVLAALAAYCLS 226
XX RESULT 78
XX AAW04571
XX ID AAW04571 standard; protein; 270 AA.
XX AC AAW04571;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 09-FEB-1997 (first entry)
XX DE HCV NS3-NS4A protease fusion.
XX KW Protease; inhibitor; surface plasmon resonance; SPR; assay;
XX KW hepatitis C virus; HCV; pNB182delta4A HT; NS3; NS4A.
XX OS Hepatitis C virus.

```

```

OS Synthetic.
OS Chimeric.
XX FH Key
XX FT Peptide
XX FT Location/Qualifiers
XX FT 5..10
XX FT /note= "hexahistidine affinity tag"
XX FT 28..209
XX FT /label= NS3
XX FT /note= "NS3 protease amino acids 1-182"
XX FT 211..261
XX FT /label= NS4A
XX FT /note= "NS4A protease amino acids 4-54"
XX PN WO9635806-A1.
XX PD 14-NOV-1996.
XX PF 09-MAY-1996; 96WO-US006385.
XX PR 12-MAY-1995; 95US-00440283.
XX PA (SCHE ) SCHERING CORP.
XX PI Taremi SS, Prorise WW;
XX DR WPI; 1996-518694/51.
XX DR N-PSDB; AAT38903.
XX PT Detecting substrate cleavage by enzyme using surface plasmon resonance -
XX PT with binding agent specific for substrate immobilised on sensor, also
XX PT used to identify enzyme, specifically protease, inhibitors.
XX PS Example 3; Page 39-40; 63pp; English.
XX CC A recombinant polypeptide (AAW04571) comprises the NS3 protease (amino
XX CC acids 1-182) of hepatitis C virus (HCV) linked to the NS4A protease
XX CC (amino acids 5-54). It does not contain an NS3/NS4A cleavage site at the
XX CC junction and is not cleaved by the autocatalytic activity of NS3. It is
XX CC encoded by cDNA construct pNB182delta4A HT (AAT38903). The fusion
XX CC protein, produced in E. coli transformants, can be used in novel high-
XX CC throughput assays for potential inhibitors of HCV protease. Surface
XX CC plasmon resonance is used to determine whether cleavage of a peptide
XX CC substrate (see also AAW04561-62) of the protease is inhibited by a test
XX CC cpd. (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 270 AA;
XX Query Match 11.0%; Score 13; DB 2; Length 270;
XX Best Local Similarity 100.0%; Pred. No. 0.0008;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAALAAYCLS 30
XX DB 214 GGVLAALAAYCLS 226
XX RESULT 79
XX AAW01645
XX ID AAW01645 standard; protein; 270 AA.
XX AC AAW01645;
XX XX
XX DT 27-AUG-2003 (revised)
XX DT 21-APR-1997 (first entry)
XX DE HCV NS3 soluble protease catalytic domain His(-V)182delta4AHT.
XX KW HCV; NS3 protease; NS4A cofactor; inhibitor; His(-V)192delta4AHT.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX OS Chimeric.

```

PH Key Region Location/Qualifiers
 FT 1. .11
 FT /label= Hexahistidine tag
 FT /note= "the hexahistidine tag facilitates affinity
 FT purification of the protein"
 FT 28. .209
 FT /label= Catalytic domain
 FT /note= "amino acid residues 28-209 correspond to amino
 FT acids 1-182 of the catalytic domain of HCV NS3 protease"
 FT 211. .261
 FT /label= NS4A
 FT /note= "amino acid residues 211-261 correspond to amino
 FT acids 4-54 of the NS4A cofactor; NS4A enhances the
 FT processing activity of the NS3 protease catalytic domain"
 FT 262. .270
 FT /label= Solubilising motif
 FT /note= "the solubilising motif neutralises the
 FT hydrophobicity of the catalytic domain, thereby improving
 FT solubility"
 XX
 XX W09636702-A2.
 XX
 XX 21-NOV-1996.
 XX
 XX 09-MAY-1996; 96WO-US006387.
 XX
 XX 12-MAY-1995; 95US-00440409.
 XX (SCHE) SCHERING CORP.
 XX
 XX Daamhapatra B, Murray MG, Ramanathan L, Butkiewicz NJ;
 XX WPI; 1997-012081/01.
 XX N-PSDB; AAT58396.
 XX
 XX Bacterially produced Hepatitis C virus NS3 protease(s) - which are
 XX denatured and re-folded to produce soluble, active enzyme.
 XX
 XX Example 2; Page 49-51; 71pp; English.
 XX
 XX His(-V)182delta4AHT (AAW01645) comprises a fusion between the catalytic
 XX domain (see also AAW01641) of hepatitis C virus (HCV) NS3 protease and
 XX the HCV NS4A polypeptide (see also AAW01651) plus a C-terminal
 XX solubilising motif (see also AAW01642) and N-terminal hexahistidine tag.
 XX A cDNA clone (AAT58396) can be used to express large quantities of the
 XX polypeptide in E. coli host cells. Soluble, active NS3 proteases (see
 XX also AAW01643-50) are useful in high throughput screens for the detection
 XX of inhibitors of the protease that can be used as therapeutic agents
 XX against HCV and also for structural studies. (Updated on 27-AUG-2003 to
 XX correct OS field.)
 XX
 XX Sequence 270 AA;
 XX
 XX Query Match 11.0%; Score 13; DB 2; Length 270;
 XX Best Local Similarity 100.0%; Pred. NO. 0.0008;
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 18 GGVLAALAAAYCLS 30
 XX |||||
 XX DB 214 GGVLAALAAAYCLS 226
 XX
 XX RESULT 80
 XX AAR30064
 XX ID AAR30064 standard; protein; 313 AA.
 XX
 XX AC AAR30064;
 XX
 XX 24-OCT-2003 (revised)
 XX 25-MAR-2003 (revised)
 XX 14-MAY-1993 (first entry)
 XX
 XX HCV NS3/NS4 non-structural region.

XX PCR; amplification; prototype; HCV pt.
 XX Hepatitis C virus; HCVel.
 XX W09221759-A1.
 XX
 XX 10-DEC-1992.
 XX
 XX 04-JUN-1992; 92WO-FR000501.
 XX
 XX 06-JUN-1991; 91FR-00006882.
 XX (INSP) INST PASTEUR.
 XX
 XX Brechot C, Krensdorf D, Porchon C;
 XX WPI; 1992-433657/52.
 XX
 XX New nucleotide and peptide sequences - specific for French isolate of
 XX hepatitis C virus and useful in diagnosing and treating related
 XX infections.
 XX
 XX Disclosure; Fig 9; 50pp; French.
 XX
 XX RNA was extracted from the serum of an HCV-positive blood donor, subjected
 XX to reverse transcription and the cDNA formed amplified by PCR.
 XX Amplification prods. were cloned, screened with a probe derived from the
 XX HCV prototype and inserts sequenced. The results showed marked
 XX conservation in the non-coding region, significant variability in the
 XX structural region (encoding envelope proteins) and reduced variability in
 XX the non-structural region. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 313 AA;
 XX
 XX Query Match 11.0%; Score 13; DB 2; Length 313;
 XX Best Local Similarity 100.0%; Pred. NO. 0.00091;
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 18 GGVLAALAAAYCLS 30
 XX |||||
 XX DB 210 GGVLAALAAAYCLS 222
 XX
 XX RESULT 81
 XX AAY22022
 XX ID AAY22022 standard; protein; 313 AA.
 XX
 XX AC AAY22022;
 XX
 XX 27-AUG-2003 (revised)
 XX 26-AUG-1999 (first entry)
 XX
 XX HCV E1 peptide sequence.
 XX
 XX HCV E1 region; monoclonal antibody; diagnosis; HCV E1-specific antigen.
 XX
 XX Hepatitis C virus.
 XX
 XX US5919454-A.
 XX
 XX 06-JUL-1999.
 XX
 XX 07-JUN-1995; 95US-00487231.
 XX
 XX 18-MAR-1993; 93US-00965285.
 XX (INSP) INST PASTEUR.
 XX
 XX Porchon C, Brechot C, Krensdorf D;
 XX WPI; 1999-394595/33.

DR N-PSDB; AAX84003.
XX Nucleotides and peptides from hepatitis C virus isolate for detecting E1-
PT specific antigens.
XX
XX
PS Claim 3; Col 17-20; 45pp; English.
XX
CC This sequence represents a hepatitis C virus (HCV) E1 region protein. The
CC invention relates to human or murine monoclonal antibodies directed
CC against a HCV E1 protein sequence. The monoclonal antibodies and their
CC fragments are useful for the in vitro diagnosis of HCV E1-specific
CC antigens. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 313 AA;
SQ

Query Match 11.0%; Score 13; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00091;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 210 GGVLAALAAAYCLS 222

RESULT 82
AAW75483
ID AAW75483 standard; protein; 313 AA.
XX
AC AAW75483;
XX
XX 27-APR-1999 (first entry)
XX
XX Hepatitis C virus NS3/NS4 protein.
XX
XX E1 region; French Hepatitis C virus; HCV; immunogen; antibody; detection;
KW immunoassay.
KW
XX Hepatitis C virus.
XX
XX US5866139-A.
XX
XX 02-FEB-1999.
XX
XX 07-JUN-1995; 95US-00483695.
XX
XX 18-MAR-1993; 93US-00965285.
XX
XX (INSP) INST PASTEUR.
XX
XX Porchon C, Kremadorf D, Brechot C;
XX
XX WPI; 1999-141865/12.
DR N-PSDB; AAX16760.
XX
XX New isolated and purified Hepatitis C virus E1 peptides - useful for
PT vaccine production or diagnostic purposes.
PT
XX Claim 21; Col 15-18; 45pp; English.
XX
XX The sequence represents the NS3/NS4 protein from a French Hepatitis C
CC virus (HCV) isolate. The protein or peptides derived from it can be: (i)
CC conjugated to a carrier protein and used as immunogens for eliciting
CC protective antibodies; or (ii) labelled, and used as immunoassay reagents
CC for detecting antibodies specific for HCV E1
XX
XX Sequence 313 AA;
SQ

Query Match 11.0%; Score 13; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00091;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 210 GGVLAALAAAYCLS 222

RESULT 83
AAR33557
ID AAR33557 standard; protein; 342 AA.
XX
AC AAR33557;
XX
XX 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
XX Antigen produced by plasmid pHCV-54.
DE
XX
XX Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
KW immunodot assay.
XX
XX Hepatitis C virus.
OS
XX W09304087-A1.
FN
XX 04-MAR-1993.
PD
XX 21-AUG-1992; 92WO-US007187.
PF
XX 21-AUG-1991; 91US-00748566.
PR
XX (ABBO) ABBOTT LAB.
PA
XX Desai SM, Casey JM, Rupprecht KR, Devare SG;
PI WPI; 1993-093940/11.
XX
XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
PT hepatitis C virus.
XX
XX Example 10; Page 41-42; 206pp; English.
PS
XX A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low
CC levels as a CKS fusion protein in E.coli. Deletion studies were performed
CC to identify the region(s) that were deleterious to the expression of the
CC HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did
CC not express the HCV CKS C100A protein. An internal deletion of 23 amino
CC acids (AAs 1575-1597) was constructed using DdeI and designated pHCV-54.
CC The deletion expressed well as a CKS fusion in E.coli. (Updated on 25-MAR
CC -2003 to correct FN field.)
XX
XX Sequence 342 AA;
SQ

Query Match 11.0%; Score 13; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 0.00098;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 320 GGVLAALAAAYCLS 332

RESULT 84
AAB69013
ID AAB69013 standard; peptide; 342 AA.
XX
AC AAB69013;
XX
XX 17-APR-2001 (first entry)
DT
XX
XX HCV recombinant antigen pHCV-54 amino acid sequence SEQ ID NO:37.
DE
XX Hepatitis C virus; HCV; antigen; detection; antibody.
KW
XX Hepatitis C virus.
OS
XX

PT	New recombinant antigens representing distinct antigenic regions of	
PT	Hepatitis C virus (HCV) genome, useful for detection of antibodies and	
PT	antigens in body fluids of individuals exposed to HCV.	
XX		
XX	Example 16; Col 159-160; 167pp; English.	
XX		
CC	The present invention describes recombinant Hepatitis C virus (HCV)	
CC	antigens (I). (I) is useful as a reagent for the detection of antibodies	
CC	and antigen in body fluids from individuals exposed to HCV. The HCV assay	
CC	uses reliable and efficient reagents and methods to accurately detect the	
CC	presence of HCV antibodies in samples obtained from individuals suspected	
CC	of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and	
CC	AAB59001 to AAB59032 represent sequences used in the exemplification of	
CC	the present invention	
XX		
XX		
SQ	Sequence 344 AA;	
	Query Match 11.0%; Score 13; DB 4; Length 344;	
	Best Local Similarity 100.0%; Pred. NO. 0.00098;	
	Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	18 GGVLAALAAAYCLS 30	
DB	322 GGVLAALAAAYCLS 334	
RESULT 88		
ABW01879		
ID	ABW01879 standard; protein; 344 AA.	
XX	AC	
XX	ABW01879;	
XX	AC	
DT	12-FEB-2004 (first entry)	
XX		
DE	HCV CKS-C100A recombinant antigen, pHCV-55.	
XX		
KW	Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;	
KW	CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;	
KW	fusion protein.	
XX		
OS	Chimeric - Hepatitis C virus.	
OS	Chimeric - Escherichia coli.	
OS	Chimeric - Unidentified.	
XX		
PN	US6593083-B1.	
XX		
PD	15-JUL-2003.	
XX		
PF	17-OCT-2000; 2000US-00690359.	
XX		
XX	24-AUG-1990; 90US-00572822.	
PR	07-NOV-1990; 90US-00614069.	
PR	21-AUG-1991; 91US-00748561.	
PR	21-AUG-1991; 91US-00748561.	
PR	21-AUG-1991; 91US-00748566.	
PR	19-NOV-1992; 92US-00989843.	
PR	10-JAN-1994; 94US-00179896.	
PR	01-MAY-1996; 96US-00646757.	
PR	02-JUN-1997; 97US-00867611.	
XX		
PA	(DEVA/) DEVARE S G.	
PA	(DESA/) DESAI S M.	
PA	(CASE/) CASEY J M.	
PA	(DAIL/) DAILEY S H.	
PA	(DAWS/) DAWSON G J.	
PA	(GUTI/) GUTIERREZ R A.	
PA	(LESN/) LESNIEWSKI R R.	
PA	(STEW/) STEWART J L.	
PA	(RUPP/) RUPPRECHT K R.	
XX		
PI	Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;	
PI	Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;	
XX		

DR WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

PT antibody is immunologically reactive with a Hepatitis C virus (HCV)

PT antigen by contacting the fluid sample, which may contain a HCV antibody

PT with at a polypeptide.

XX Example 16; Col 157-160; 168pp; English.

XX The invention relates to a method for identifying the presence of an

CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.

CC The method involves providing a fluid sample containing at least one HCV

CC antibody, contacting the fluid sample with at least one polypeptide or

CC recombinant fusion protein for complexing the antibody with the

CC polypeptide or recombinant fusion protein to provide an antibody-

CC polypeptide complex and detecting the complex. The present sequence is

CC pHCV-55 fusion protein which comprises Escherichia coli CKS (Ctp: CMP-3-

CC deoxy-manno-octuloseonate cytidyl) transferase or CMP-KDO synthase)

CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence

CC is used in the invention

XX

SQ Sequence 344 AA;

Query Match 11.0%; Score 13; DB 7; Length 344;

Best Local Similarity 100.0%; Pred. No. 0.00098;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30

Db 322 GGVLAALAAAYCLS 334

|||||

RESULT 89

AAAR33559

ID AAR33559 standard; protein; 352 AA.

XX

AC AAR33559;

DT 25-MAR-2003 (revised)

DT 01-JUL-1993 (first entry)

XX

DE Antigen produced by plasmid pHCV-94.

XX Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;

KW immunodot assay.

XX

OS Hepatitis C virus.

XX

PN WO9304087-A1.

XX

PD 04-MAR-1993.

XX

PF 21-AUG-1992; 92WO-US007187.

XX

PR 21-AUG-1991; 91US-00748566.

XX

PA (ABBO) ABBOTT LAB.

XX

PI Desai SM, Casey JM, Rupprecht KR, Devare SG;

XX

DR WPI; 1993-093940/11.

XX

PT Hepatitis C assay using recombinant C-100 region antigens - for detecting

PT antibodies and antigen in body fluids from individuals exposed to

PT hepatitis C virus.

XX

XX Example 10; Page 43-45; 206pp; English.

XX A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low

CC levels as a CKS fusion protein in E.coli. Deletion studies were performed

CC to identify the region(s) that were deleterious to the expression of the

CC HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did

CC not express the HCV CKS C100A protein. An internal deletion of 21 amino

CC

CC acids (AAs 1600-1620) was constructed using NlaIV/HaeIII and designated

CC pHCV-55 (see AAR33557). The deletion expressed well as a CKS fusion in

CC E.coli. The amino acids deleted in pHCV-55 were sequentially replaced

CC from the carboxy-terminal end using a fragment replacement method.

CC Plasmid pHCV-94 is one of the constructs produced by this process.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 352 AA;

Query Match 11.0%; Score 13; DB 2; Length 352;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30

Db 330 GGVLAALAAAYCLS 342

|||||

RESULT 90

AAAB69015

ID AAB69015 standard; peptide; 352 AA.

XX

AC AAB69015;

DT 17-APR-2001 (first entry)

XX

DE HCV recombinant antigen pHCV-94 amino acid sequence SEQ ID NO:39.

XX Hepatitis C virus; HCV; antigen; detection; antibody.

XX

OS Hepatitis C virus.

XX

PN US6172189-B1.

XX

PD 09-JAN-2001.

XX

PF 02-JUN-1997; 97US-00867611.

XX

PR 24-AUG-1990; 90US-00572822.

PR 07-NOV-1990; 90US-00614069.

PR 21-AUG-1991; 91US-00748561.

PR 21-AUG-1991; 91US-00748565.

PR 21-AUG-1991; 91US-00748566.

PR 19-NOV-1992; 92US-00989843.

PR 10-JAN-1994; 94US-00179896.

PR 01-MAY-1996; 96US-00646757.

XX

PA (ABBO) ABBOTT LAB.

XX

PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

PI Gutierrez RA, Lesniowski RR, Stewart JL, Rupprecht KR;

XX

DR WPI; 2001-122352/13.

XX

PT New recombinant antigens representing distinct antigenic regions of

PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and

PT antigens in body fluids of individuals exposed to HCV.

XX

XX Example 16; Col 161-162; 167pp; English.

XX The present invention describes recombinant Hepatitis C virus (HCV)

CC antigens (I). (I) is useful as a reagent for the detection of antibodies

CC and antigen in body fluids from individuals exposed to HCV. The HCV assay

CC uses reliable and efficient reagents and methods to accurately detect the

CC presence of HCV antibodies in samples obtained from individuals suspected

CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and

CC AAB69001 to AAB69032 represent sequences used in the exemplification of

CC the present invention

XX

SQ Sequence 352 AA;

Query Match 11.0%; Score 13; DB 4; Length 352;

Best Local Similarity 100.0%; Pred. No. 0.001;

```
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
Db 330 GGVLAALAAAYCLS 342
|||||
RESULT 91
ABW01880
ID ABW01880 standard; protein; 352 AA.
XX AC
XX ABW01880;
XX
XX 12-FEB-2004 (first entry)
XX
XX HCV CKS-C100A recombinant antigen, pHCV-94.
XX
XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
KW fusion protein.
XX
XX Chimeric - Hepatitis C virus.
OS
XX Chimeric - Escherichia coli.
OS
XX Chimeric - Unidentified.
XX
XX US6593083-B1.
XX
XX 15-JUL-2003.
XX
XX 17-OCT-2000; 2000US-00690359.
XX
XX 24-AUG-1990; 90US-00572822.
XX
XX 07-NOV-1990; 90US-00614069.
XX
XX 21-AUG-1991; 91US-00748561.
XX
XX 21-AUG-1991; 91US-00748565.
XX
XX 21-AUG-1991; 91US-00748566.
XX
XX 19-NOV-1992; 92US-00989843.
XX
XX 10-JAN-1994; 94US-00179896.
XX
XX 01-MAY-1996; 96US-00646757.
XX
XX 02-JUN-1997; 97US-00867611.
XX
XX (DEVA/) DESVARE S G.
XX
XX (DESA/) DESAI S M.
XX
XX (CASE/) CASEY J M.
XX
XX (DAIL/) DAILEY S H.
XX
XX (DAWS/) DAWSON G J.
XX
XX (GUTI/) GUTIERREZ R A.
XX
XX (LESN/) LESNIEWSKI R R.
XX
XX (STEW/) STEWART J L.
XX
XX (RUPP/) RUPPRECHT K R.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX
XX WPI; 2003-828264/77.
XX
XX Identifying the presence of an antibody in a fluid sample, where the
PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
PT antigen by contacting the fluid sample, which may contain a HCV antibody
PT with at a polypeptide.
XX
XX Example 16; Col 159-162; 168pp; English.
XX
XX The invention relates to a method for identifying the presence of an
CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
CC The method involves providing a fluid sample containing at least one HCV
CC antibody, contacting the fluid sample with at least one polypeptide or
CC recombinant fusion protein for complexing the antibody with the
CC polypeptide or recombinant fusion protein to provide an antibody-
CC polypeptide complex and detecting the complex. The present sequence is
CC pHCV-94 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
```

CC is used in the invention
XX
SQ Sequence 352 AA;

Query Match 11.0%; Score 13; DB 7; Length 352;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 330 GGVLAALAAAYCLS 342
|||||

RESULT 92
AAR33560
ID AAR33560 standard; protein; 357 AA.

XX AC AAR33560;

XX DT 25-MAR-2003 (revised)
XX DT 01-JUL-1993 (first entry)

XX DE Antigen produced by plasmid pHCV-95.

XX KW Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
XX KW immunodot assay.

XX OS Hepatitis C virus.

XX PN WO9304087-A1.

XX PD 04-MAR-1993.

XX PF 21-AUG-1992; 92WO-US007187.

XX PR 21-AUG-1991; 91US-00748566.

XX PA (ABBO) ABBOTT LAB.

XX PI Desai SM, Casey JM, Rupprecht KR, Devare SG;

XX WPI; 1993-093940/11.

XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.

XX Example 10; Page 45-46; 206pp; English.

XX A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low
XX levels as a CKS fusion protein in E. coli. Deletion studies were performed
XX to identify the region(s) that were deleterious to the expression of the
XX HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did
XX not express the HCV CKS C100A protein. An internal deletion of 21 amino
XX acids (AAs 1600-1620) was constructed using NlaIV/HaeIII and designated
XX pHCV-55 (see AAR33557). The deletion expressed well as a CKS fusion in
XX E. coli. The amino acids deleted in pHCV-55 were sequentially replaced
XX from the carboxy-terminal end using a fragment replacement method.
XX Plasmid pHCV-95 is one of the constructs produced by this process.
XX (Updated on 25-MAR-2003 to correct FN field.)

XX SQ Sequence 357 AA;

Query Match 11.0%; Score 13; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 335 GGVLAALAAAYCLS 347
|||||

RESULT 93

```
AAB69016
ID AAB69016 standard; peptide; 357 AA.
XX
AC AAB69016;
XX
DT 17-APR-2001 (first entry)
XX
DE HCV recombinant antigen pHCV-95 amino acid sequence SEQ ID NO:40.
XX
KW Hepatitis C virus; HCV; antigen; detection; antibody.
XX
OS Hepatitis C virus.
XX
PN US6172189-B1.
XX
PD 09-JAN-2001.
XX
PF 02-JUN-1997; 97US-00867611.
XX
PR 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX
DR WPI; 2001-122352/13.
XX
PT New recombinant antigens representing distinct antigenic regions of
PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
PT antigens in body fluids of individuals exposed to HCV.
XX
PS Example 16; Col 163-164; 167pp; English.
XX
CC The present invention describes recombinant Hepatitis C virus (HCV)
CC antigens (I). (I) is useful as a reagent for the detection of antibodies
CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
CC uses reliable and efficient reagents and methods to accurately detect the
CC presence of HCV antibodies in samples obtained from individuals suspected
CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
CC AAB69001 to AAB69032 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 357 AA;
Query Match 11.0%; Score 13; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 18 GGVLAAALAAAYCLS 30
Db 335 GGVLAAALAAAYCLS 347
RESULT 94
ABW01881
ID ABW01881 standard; protein; 357 AA.
XX
AC ABW01881;
XX
DT 12-FEB-2004 (first entry)
XX
DE HCV CKS-C100A recombinant antigen, pHCV-95.
XX
KW Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
```

```
fusion protein.
KW
XX Chimeric - Hepatitis C virus.
OS Chimeric - Escherichia coli.
OS Chimeric - Unidentified.
XX
PN US6593083-B1.
XX
PD 15-JUL-2003.
XX
PF 17-OCT-2000; 2000US-00690359.
XX
PR 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
PR 02-JUN-1997; 97US-00867611.
XX
PA (DEVA/) DEVARE S G.
PA (DESA/) DESAI S M.
PA (CASE/) CASEY J M.
PA (DAIL/) DAILEY S H.
PA (DAWS/) DAWSON G J.
PA (GUTI/) GUTIERREZ R A.
PA (LESN/) LESNIEWSKI R R.
PA (STEW/) STEWART J L.
PA (RUPP/) RUPPRECHT K R.
XX
PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX
DR WPI; 2003-828264/77.
XX
PT Identifying the presence of an antibody in a fluid sample, where the
PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
PT antigen by contacting the fluid sample, which may contain a HCV antibody
PT with at a polypeptide.
XX
PS Example 16; Col 161-164; 168pp; English.
XX
CC The invention relates to a method for identifying the presence of an
CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
CC The method involves providing a fluid sample containing at least one HCV
CC antibody, contacting the fluid sample with at least one polypeptide or
CC recombinant fusion protein for complexing the antibody with the
CC polypeptide or recombinant fusion protein to provide an antibody-
CC polypeptide complex and detecting the complex. The present sequence is
CC pHCV-95 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
CC is used in the invention
XX
SQ Sequence 357 AA;
Query Match 11.0%; Score 13; DB 7; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 18 GGVLAAALAAAYCLS 30
Db 335 GGVLAAALAAAYCLS 347
RESULT 95
AAR33561
ID AAR33561 standard; protein; 362 AA.
XX
AC AAR33561;
XX
```

DT 25-MAR-2003 (revised)
 DT 01-JUL-1993 (first entry)
 XX
 DE Antigen produced by plasmid pHCV-96.
 XX
 KW Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
 KW immunodot assay.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9304087-A1.
 XX
 XX 04-MAR-1993.
 PD
 XX 21-AUG-1992; 92WO-US007187.
 PP
 XX 21-AUG-1991; 91US-00748566.
 PR
 XX (ABBO) ABBOTT LAB.
 PA
 XX Desai SM, Casey JM, Rupprecht KR, Devare SG;
 PI WPI; 1993-093940/11.
 XX
 DR Hepatitis C assay using recombinant C-100 region antigens - for detecting
 XX antibodies and antigen in body fluids from individuals exposed to
 PT hepatitis C virus.
 PT
 XX Example 10; Page 46-48; 206pp; English.
 PS
 XX A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low
 CC levels as a CKS fusion protein in E.coli. Deletion studies were performed
 CC to identify the region(s) that were deleterious to the expression of the
 CC HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did
 CC not express the HCV CKS C100A protein. An internal deletion of 21 amino
 CC acids (AAs 1600-1620) was constructed using NlaIV/HaeIII and designated
 CC pHCV-55 (see AAR33557). The deletion expressed well as a CKS fusion in
 CC E.coli. The amino acids deleted in pHCV-55 were sequentially replaced
 CC from the carboxy-terminal end using a fragment replacement method.
 CC Plasmid pHCV-96 is one of the constructs produced by this process.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 362 AA;
 SQ
 Query Match 11.0%; Score 13; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAAALAAAYCLS 30
 Db 340 GGVLAAALAAAYCLS 352
 RESULT 96
 ID AAB69017 standard; peptide; 362 AA.
 AC AAB69017;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE HCV recombinant antigen pHCV-96 amino acid sequence SEQ ID NO:41.
 XX
 KW Hepatitis C virus; HCV; antigen; detection; antibody.
 XX
 OS Hepatitis C virus.
 XX
 PN US6172189-B1.
 XX
 XX 09-JAN-2001.
 PD
 XX 02-JUN-1997; 97US-00867611.
 PF
 XX

PR 24-AUG-1990; 90US-00572822.
 PR 07-NOV-1990; 90US-00614069.
 PR 21-AUG-1991; 91US-00748561.
 PR 21-AUG-1991; 91US-00748565.
 PR 21-AUG-1991; 91US-00748566.
 PR 19-NOV-1992; 92US-00989843.
 PR 10-JAN-1994; 94US-00179896.
 PR 01-MAY-1996; 96US-00646757.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 XX WPI; 2001-122352/13.
 DR
 XX New recombinant antigens representing distinct antigenic regions of
 PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
 PT antigens in body fluids of individuals exposed to HCV.
 XX
 XX Example 16; Col 165-166; 167pp; English.
 PS
 XX The present invention describes recombinant Hepatitis C virus (HCV)
 CC antigens (I). (I) is useful as a reagent for the detection of antibodies
 CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
 CC uses reliable and efficient reagents and methods to accurately detect the
 CC presence of HCV antibodies in samples obtained from individuals suspected
 CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
 CC AAB69001 to AAB69032 represent sequences used in the exemplification of
 CC the present invention
 XX
 XX Sequence 362 AA;
 SQ
 Query Match 11.0%; Score 13; DB 4; Length 362;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAAALAAAYCLS 30
 Db 340 GGVLAAALAAAYCLS 352
 RESULT 97
 ID AABW01882
 XX AABW01882 standard; protein; 362 AA.
 ID
 AC AABW01882;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE HCV CKS-C100A recombinant antigen, pHCV-96.
 XX
 KW Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
 KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
 KW fusion protein.
 XX
 OS Chimeric - Hepatitis C virus.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Unidentified.
 XX
 PN US6593083-B1.
 XX
 XX 15-JUL-2003.
 PD
 XX 17-OCT-2000; 2000US-00690359.
 PF
 XX 24-AUG-1990; 90US-00572822.
 PR 07-NOV-1990; 90US-00614069.
 PR 21-AUG-1991; 91US-00748561.
 PR 21-AUG-1991; 91US-00748565.
 PR 21-AUG-1991; 91US-00748566.
 PR 19-NOV-1992; 92US-00989843.
 PR 10-JAN-1994; 94US-00179896.
 PR

CC A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low
 CC levels as a CKS fusion protein in E.coli. Deletion studies were performed.
 CC to identify the region(s) that were deleterious to the expression of the
 CC HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did
 CC not express the HCV CKS C100A protein. An internal deletion of 21 amino
 CC acids (AAs 1600-1620) was constructed using NlaIV/HaeIII and designated
 CC pHCV-55 (see AAR33557). The deletion expressed well as a CKS fusion in
 CC E.coli. The amino acids deleted in pHCV-55 were sequentially replaced
 CC from the carboxy-terminal end using a fragment replacement method.
 CC Plasmid pHCV-202 is one of the constructs produced by this process.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX Sequence 363 AA;

SQ

Query Match 11.0%; Score 13; DB 2; Length 363;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

341 GGVLAALAAAYCLS 353

RESULT 100

AAR90933
 ID AAR90933 standard; protein; 363 AA.

XX
 AC AAR90933;

XX
 DT 25-MAR-2003 (revised)

XX
 DT 15-MAY-1996 (first entry)

XX
 DE HCV antigen C100.

XX
 DE Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;

XX
 KW antibodies.

XX
 OS Hepatitis C virus.

XX
 PN EP693687-A1.

XX
 PD 24-JAN-1996.

XX
 PF 03-APR-1991; 95EP-00114016.

XX
 PR 04-APR-1990; 90US-00504352.

XX
 PA (CHIR) CHIRON CORP.

XX
 PI Houghton M, Choo Q, Kuo G;

XX
 DR WPI; 1996-117956/13.

XX
 PT Combinations of synthetic Hepatitis C Virus antigens - provide more

XX
 PT effective diagnosis of Non-A, Non-B Hepatitis.

XX
 PS Claim 6; Fig 1(A-Y); 53pp; English.

XX
 CC The combination comprises an HCV antigen from the C domain (pref. C22 -

XX
 CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -

XX
 CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5

XX
 CC (AAR90934) domain. The antigens may in the form of a fusion protein, a

XX
 CC simple physical mixture, or the individual antigens commonly bound to a

XX
 CC solid matrix. They are pref. prepd. by recombinant DNA techniques

XX
 CC (primers are given in AAR12711-R12716), but can be synthesised or

XX
 CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003

XX
 CC to correct PF field.)

XX
 SQ Sequence 363 AA;

Query Match 11.0%; Score 13; DB 2; Length 363;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

18 GGVLAALAAAYCLS 30

|||||

341 GGVLAALAAAYCLS 353

RESULT 102

ABW01884

ID ABW01884 standard; protein; 363 AA.

XX
 AC ABW01884;

QY 18 GGVLAALAAAYCLS 30

|||||

96 GGVLAALAAAYCLS 108

RESULT 101

AAB69019

ID AAB69019 standard; peptide; 363 AA.

XX
 AC AAB69019;

XX
 DT 17-APR-2001 (first entry)

XX
 DE HCV recombinant antigen pHCV-202 amino acid sequence SEQ ID NO:43.

XX
 KW Hepatitis C virus; HCV; antigen; detection; antibody.

XX
 OS Hepatitis C virus.

XX
 PN US6172189-B1.

XX
 PD 09-JAN-2001.

XX
 PF 02-JUN-1997; 97US-00867611.

XX
 PR 24-AUG-1990; 90US-00572822.

XX
 PR 07-NOV-1990; 90US-00614069.

XX
 PR 21-AUG-1991; 91US-00748561.

XX
 PR 21-AUG-1991; 91US-00748565.

XX
 PR 19-NOV-1992; 92US-00989843.

XX
 PR 10-JAN-1994; 94US-00179896.

XX
 PR 01-MAY-1996; 96US-00646757.

XX
 PA (ABBO) ABBOTT LAB.

XX
 PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX
 DR WPI; 2001-122352/13.

XX
 PT New recombinant antigens representing distinct antigenic regions of

XX
 PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and

XX
 PT antigens in body fluids of individuals exposed to HCV.

XX
 PS Example 16; Col 169-172; 167pp; English.

XX
 CC The present invention describes recombinant Hepatitis C virus (HCV)

XX
 CC antigens (I). (I) is useful as a reagent for the detection of antibodies

XX
 CC and antigen in body fluids from individuals exposed to HCV. The HCV assay

XX
 CC uses reliable and efficient reagents and methods to accurately detect the

XX
 CC presence of HCV antibodies in samples obtained from individuals suspected

XX
 CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and

XX
 CC AAB69001 to AAB69032 represent sequences used in the exemplification of

XX
 CC the present invention

XX
 SQ Sequence 363 AA;

Query Match 11.0%; Score 13; DB 4; Length 363;

Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

341 GGVLAALAAAYCLS 353

RESULT 102

ABW01884

ID ABW01884 standard; protein; 363 AA.

XX
 AC ABW01884;

```
XX 12-FEB-2004 (first entry)
DT HCV CKS-Cl00A recombinant antigen, pHCV-202.
XX
DE Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
XX fusion protein.
XX
OS Chimeric - Hepatitis C virus.
OS Chimeric - Escherichia coli.
OS Chimeric - Unidentified.
XX
PN US6593083-B1.
XX
XX 15-JUL-2003.
XX
XX 17-OCT-2000; 2000US-00690359.
XX
XX 24-AUG-1990; 90US-00572822.
XX 07-NOV-1990; 90US-00614069.
XX 21-AUG-1991; 91US-00748561.
XX 21-AUG-1991; 91US-00748565.
XX 21-AUG-1991; 91US-00748566.
XX 19-NOV-1992; 92US-00989843.
XX 10-JAN-1994; 94US-00179896.
XX 01-MAY-1996; 96US-00846757.
XX 02-JUN-1997; 97US-00867611.
XX
XX (DEVA/) DEVARE S G.
XX (DESA/) DESAI S M.
XX (CASE/) CASEY J M.
XX (DAIL/) DAILEY S H.
XX (DAWS/) DAWSON G J.
XX (GUTI/) GUTIERREZ R A.
XX (LESN/) LESNIEWSKI R R.
XX (STEW/) STEWART J L.
XX (RUPP/) RUPPRECHT K R.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2003-828264/77.
XX
XX Identifying the presence of an antibody in a fluid sample, where the
XX antibody is immunologically reactive with a Hepatitis C virus (HCV)
XX antigen by contacting the fluid sample, which may contain a HCV antibody
XX with at a polypeptide.
XX
XX Example 16; Col 169-170; 168pp; English.
XX
XX The invention relates to a method for identifying the presence of an
XX antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
XX The method involves providing a fluid sample containing at least one HCV
XX antibody, contacting the fluid sample with at least one polypeptide or
XX recombinant fusion protein for complexing the antibody with the
XX polypeptide or recombinant fusion protein to provide an antibody-
XX polypeptide complex and detecting the complex. The present sequence is
XX pHCV-202 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
XX deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
XX enzyme, linker and HCV (non- structural region) NS4 region. This sequence
XX is used in the invention
XX
XX Sequence 363 AA;
XX
XX Query Match 11.0%; Score 13; DB 7; Length 363;
XX Best Local Similarity 100.0%; Pred. No. 0.001;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 341 GGVLAALAAAYCLS 353
```

```
RESULT 103
AAR33564
ID AAR33564 standard; protein; 364 AA.
XX
XX AAR33564;
AC
XX 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
XX Antigen produced by plasmid pHCV-203.
XX
XX Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
XX immunodot assay.
XX
XX Hepatitis C virus.
XX
XX WO9304087-A1.
XX
XX 04-MAR-1993.
XX
XX 21-AUG-1992; 92WO-US007187.
XX
XX 21-AUG-1991; 91US-00748566.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Desai SM, Casey JM, Rupprecht KR, Devare SG;
XX WPI; 1993-093940/11.
XX
XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.
XX
XX Example 10; Page 51-52; 206pp; English.
XX
XX A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low
XX levels as a CKS fusion protein in E.coli. Deletion studies were performed
XX to identify the region(s) that were deleterious to the expression of the
XX HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did
XX not express the HCV CKS C100A protein. An internal deletion of 21 amino
XX acids (AAs 1600-1620) was constructed using NlaIV/HaeIII and designated
XX pHCV-55 (see AAR33557). The deletion expressed well as a CKS fusion in
XX E.coli. The amino acids deleted in pHCV-55 were sequentially replaced
XX from the carboxy-terminal end using a fragment replacement method.
XX Plasmid pHCV-203 is one of the constructs produced by this process.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 364 AA;
XX
XX Query Match 11.0%; Score 13; DB 2; Length 364;
XX Best Local Similarity 100.0%; Pred. No. 0.001;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 342 GGVLAALAAAYCLS 354
XX
XX RESULT 104
AAB69020
ID AAB69020 standard; peptide; 364 AA.
XX
XX AAB69020;
AC
XX 17-APR-2001 (first entry)
DT
XX HCV recombinant antigen pHCV-203 amino acid sequence SEQ ID NO:44.
XX
XX Hepatitis C virus; HCV; antigen; detection; antibody.
XX
XX Hepatitis C virus.
XX
```

```

XX US6172189-B1.
XX
XX 09-JAN-2001.
XX
XX 02-JUN-1997; 97US-00867611.
XX
XX 24-AUG-1990; 90US-00572822.
XX 07-NOV-1990; 90US-00614069.
XX 21-AUG-1991; 91US-00748561.
XX 21-AUG-1991; 91US-00748565.
XX 21-AUG-1991; 91US-00748566.
XX 19-NOV-1992; 92US-00989843.
XX 10-JAN-1994; 94US-00179896.
XX 01-MAY-1996; 96US-00646757.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2001-122352/13.
XX
XX New recombinant antigens representing distinct antigenic regions of
XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and
XX antigens in body fluids of individuals exposed to HCV.
XX
XX Example 16; Col 171-174; 167pp; English.
XX
XX The present invention describes recombinant Hepatitis C virus (HCV)
XX antigens (I). (I) is useful as a reagent for the detection of antibodies
XX and antigen in body fluids from individuals exposed to HCV. The HCV assay
XX uses reliable and efficient reagents and methods to accurately detect the
XX presence of HCV antibodies in samples obtained from individuals suspected
XX of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
XX AAB69001 to AAB69032 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 364 AA;
XX
XX Query Match 11.0%; Score 13; DB 4; Length 364;
XX Best Local Similarity 100.0%; Pred. No. 0.001;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 342 GGVLAALAAAYCLS 354
XX
XX RESULT 105
XX AABW01885
XX ID AABW01885 standard; protein; 364 AA.
XX
XX AC AABW01885;
XX
XX 12-FEB-2004 (first entry)
XX
XX HCV CKS-C100A recombinant antigen, pHCV-203.
XX
XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
XX CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
XX fusion protein.
XX
XX Chimeric - Hepatitis C virus.
XX OS Chimeric - Escherichia coli.
XX OS Chimeric - Unidentified.
XX
XX US6593083-B1.
XX
XX 15-JUL-2003.
XX
XX 17-OCT-2000; 2000US-00690359.
XX

```

```

PR 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
PR 02-JUN-1997; 97US-00867611.
XX
XX (DEVA/) DEVARE S G.
XX (DESA/) DESAI S M.
XX (CASE/) CASEY J M.
XX (DAIL/) DAILEY S H.
XX (DAWS/) DAWSON G J.
XX (GUTI/) GUTIERREZ R A.
XX (LESN/) LESNIEWSKI R R.
XX (STEW/) STEWART J L.
XX (RUPP/) RUPPRECHT K R.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2003-828264/77.
XX
XX Identifying the presence of an antibody in a fluid sample, where the
XX antibody is immunologically reactive with a Hepatitis C virus (HCV)
XX antigen by contacting the fluid sample, which may contain a HCV antibody
XX with at a polypeptide.
XX
XX Example 16; Col 171-172; 168pp; English.
XX
XX The invention relates to a method for identifying the presence of an
XX antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
XX The method involves providing a fluid sample containing at least one HCV
XX antibody, contacting the fluid sample with at least one polypeptide or
XX recombinant fusion protein for complexing the antibody with the
XX polypeptide or recombinant fusion protein to provide an antibody-
XX polypeptide complex and detecting the complex. The present sequence is
XX pHCV-203 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
XX deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
XX enzyme, linker and HCV (non- structural region) NS4 region. This sequence
XX is used in the invention
XX
XX Sequence 364 AA;
XX
XX Query Match 11.0%; Score 13; DB 7; Length 364;
XX Best Local Similarity 100.0%; Pred. No. 0.001;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 342 GGVLAALAAAYCLS 354
XX
XX RESULT 106
XX AAR33562
XX ID AAR33562 standard; protein; 365 AA.
XX
XX AC AAR33562;
XX
XX 25-MAR-2003 (revised)
XX 01-JUL-1993 (first entry)
XX
XX Antigen produced by plasmid pHCV-97.
XX
XX Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
XX immunodot assay.
XX
XX Hepatitis C virus.
XX
XX W09304087-A1.
XX

```

```

PD 04-MAR-1993.
XX
PF 21-AUG-1992; 92WO-US007187.
XX
PR 21-AUG-1991; 91US-00748566.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Desai SM, Casey JM, Rupprecht KR, Devare SG;
XX WPI; 1993-093940/11.
XX
DR Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.
XX
PS Example 10; Page 48-49; 206pp; English.
XX
CC A synthetic gene encoding HCV amino acids 1569-1931 was expressed at low
XX levels as a CKS fusion protein in E.coli. Deletion studies were performed
XX to identify the region(s) that were deleterious to the expression of the
XX HCV CKS-C100 antigen. Clone pHCV-19 contained codons 1569-1677 and did
XX not express the HCV CKS C100A protein. An internal deletion of 21 amino
XX acids (Aaa 1600-1620) was constructed using NlaIV/HaeIII and designated
XX pHCV-55 (see AAR33557). The deletion expressed well as a CKS fusion in
XX E.coli. The amino acids deleted in pHCV-55 were sequentially replaced
XX from the carboxy-terminal end using a fragment replacement method.
XX Plasmid pHCV-97 is one of the constructs produced by this process.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 365 AA;

Query Match 11.0%; Score 13; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 343 GGVLAALAAAYCLS 355

RESULT 107
AAB69018
ID AAB69018 standard; peptide; 365 AA.
AC AAB69018;
XX
XX 17-APR-2001 (first entry)
DT
DE HCV recombinant antigen pHCV-97 amino acid sequence SEQ ID NO:42.
XX
XX Hepatitis C virus; HCV; antigen; detection; antibody.
XX
XX Hepatitis C virus.
XX
XX US6172189-B1.
PN
XX 09-JAN-2001.
PD
XX 02-JUN-1997; 97US-00867611.
PF
XX 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
PR 01-MAY-1996; 96US-00646757.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI

```

```

PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX
DR WPI; 2001-122352/13.
XX
XX New recombinant antigens representing distinct antigenic regions of
PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
PT antigens in body fluids of individuals exposed to HCV.
XX
XX Example 16; Col 167-170; 167pp; English.
XX
XX The present invention describes recombinant Hepatitis C virus (HCV)
CC antigens (I). (I) is useful as a reagent for the detection of antibodies
CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
CC uses reliable and efficient reagents and methods to accurately detect the
CC presence of HCV antibodies in samples obtained from individuals suspected
CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
CC AAB69001 to AAB69032 represent sequences used in the exemplification of
CC the present invention
XX
XX Sequence 365 AA;

Query Match 11.0%; Score 13; DB 4; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 343 GGVLAALAAAYCLS 355

RESULT 108
ABW01883
ID ABW01883 standard; protein; 365 AA.
XX
XX AC ABW01883;
XX
XX 12-FEB-2004 (first entry)
DT
XX
DE HCV CKS-C100A recombinant antigen, pHCV-97.
XX
XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
KW fusion protein.
XX
XX Chimeric - Hepatitis C virus.
OS Chimeric - Escherichia coli.
OS Chimeric - Unidentified.
XX
XX US6593083-B1.
PN
XX 15-JUL-2003.
PD
XX 17-OCT-2000; 2000US-00690359.
PF
XX 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
PR 02-JUN-1997; 97US-00867611.
XX
XX (DEVA/) DEVARE S G.
PA (DESA/) DESAI S M.
PA (CASE/) CASEY J M.
PA (DAIL/) DAILEY S H.
PA (DAWS/) DAWSON G J.
PA (GUTI/) GUTIERREZ R A.
PA (LESN/) LESNIEWSKI R R.
PA (STEW/) STEWART J L.
PA (RUPP/) RUPPRECHT K R.

```

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 XX WPI; 2003-828264/77.
 XX Identifying the presence of an antibody in a fluid sample, where the
 PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
 PT antigen by contacting the fluid sample, which may contain a HCV antibody
 PT with at a polypeptide.
 XX Example 16; Col 167-168; 168pp; English.
 XX The invention relates to a method for identifying the presence of an
 CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
 CC The method involves providing a fluid sample containing at least one HCV
 CC antibody, contacting the fluid sample with at least one polypeptide or
 CC recombinant fusion protein for complexing the antibody with the
 CC polypeptide or recombinant fusion protein to provide an antibody-
 CC polypeptide complex and detecting the complex. The present sequence is
 CC pHCV-97 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
 CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
 CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
 CC is used in the invention
 XX Sequence 365 AA;
 SQ Query Match 11.0%; Score 13; DB 7; Length 365;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAALAAAYCLIS 30
 Db 343 GGVLAALAAAYCLIS 355
 RESULT 109
 AAP92048
 ID AAP92048 standard; protein; 382 AA.
 XX AAP92048;
 XX 09-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 21-JAN-1991 (first entry)
 XX Carboxy-terminus of the fusion polypeptide C100-3 of HCV.
 XX Non-A non-B hepatitis; probe; vaccine; diagnosis; passive immunotherapy;
 KW antigen.
 XX Hepatitis C virus.
 OS Unidentified.
 XX Key Location/Qualifiers
 FH Region 1..27
 FT /label= SOD
 FT Region 28..42
 FT /label= adaptor
 FT Region 43..377
 FT /label= NANB
 FT Region 378..382
 FT /label= extra
 XX EP118216-A.
 XX 31-MAY-1989.
 XX 18-NOV-1988; 88EP-00310922.
 XX 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 26-FEB-1988; 88US-00161072.

PR 06-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX Houghton M, Choo QL, Kuo G;
 DR WPI; 1989-159274/22.
 DR N-PSDB; AAN92104.
 XX Purified hepatitis C virus - and associated nucleic acids and
 PT polypeptide(s).
 XX Example; Fig 36-1 - 36-2; 139pp; English.
 CC Purified hepatitis C virus (HCV) and purified or recombinant HCV nucleic
 CC acids (NAs), encoding HCV polynucleotides or epitopes, and polypeptides
 CC are claimed. HCV is a causative agent of non-A, non-B hepatitis (NANBH).
 CC The NAs may be used to design probes for detn. of HCV NAs in samples. The
 CC polypeptides may be used as immunoassay reagents and vaccines, and to
 CC produce antibodies useful for diagnosis and passive immunotherapy. The
 CC purified virus may also be used in vaccines. (Updated on 25-MAR-2003 to
 CC correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
 CC Revised record issued on 09-SEP-2004 : Correction to DE line
 XX Sequence 382 AA;
 SQ Query Match 11.0%; Score 13; DB 1; Length 382;
 Best Local Similarity 100.0%; Pred. No. 0.001;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAALAAAYCLIS 30
 Db 110 GGVLAALAAAYCLIS 122
 RESULT 110
 AAP90182
 ID AAP90182 standard; protein; 382 AA.
 XX AAP90182;
 XX 25-MAR-2003 (revised)
 DT 01-NOV-1989 (first entry)
 XX C terminus of superoxide dismutase C100 fusion polypeptide.
 XX Hepatitis C virus; fusion polypeptide; probe; vaccine;
 KW non-A,non-B hepatitis.
 XX Pan troglodytes.
 OS Key Location/Qualifiers
 FH Region 1..9
 FT Region 10..14
 FT Region 15..377
 FT Region 378..382
 XX GB2212511-A.
 XX 26-JUL-1989.
 XX 18-NOV-1988; 88GB-00027024.
 PF 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 26-OCT-1988; 88US-00263584.
 XX (CHIR) CHIRON CORP.
 PA

XX
PI Houghton M, Choo QL, Kuo G;
XX WPI; 1989-215054/30.
XX N-PSDB; AAN90334.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
XX Disclosure; Fig 36; 30pp; English.
XX
CC The sequence is of the C-terminus of the superoxide dismutase (SOD)-C100
CC fusion polypeptide (see AAN90334). The sequence encodes antigens which
CC react with antibodies in patients with non-A non-B hepatitis (NANBH). The
CC polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies
CC for immunoassay or treatment, or to produce vaccines. The regions show
CC the SOD C-terminal, an adaptor, the NANBH polypeptide, and an extra
CC sequence resp. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 382 AA;

Query Match 11.0%; Score 13; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 110 GGVLAALAAAYCLS 122
|||||

RESULT 111
AAP92024
ID AAP92024 standard; protein; 460 AA.
AC AAP92024;
XX
XX 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Polypeptide from the combined ORFs of clones 35, 36 81 and 32 of HCV.
XX
XX Hepatitis C virus (HCV); non-A, non-B hepatitis (NANBH).
XX
XX Hepatitis C virus.
OS Unidentified.
XX
XX EP318216-A.
XX
XX 31-MAY-1989.
XX
XX 18-NOV-1988; 88EP-00310922.
XX
XX 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
XX (CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
XX
XX WPI; 1989-159274/22.
DR N-PSDB; AAN92080.
XX
XX Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
XX
XX Claim 13; Fig 9; 139pp; English.

XX
CC It is the polypeptide sequence encoded in the combined open reading
CC frames of the Hepatitis C virus (HCV) cDNA of clones 35, 36, 81 and 32.
CC It is an epitope, portions of which could be used as immunoassay reagents
CC and vaccines and to generate antibodies useful in diagnosis and passive
CC immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-
CC MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI
CC field.)
XX
XX Revised record issued on 09-SEP-2004 : Correction to DE line
XX
SQ Sequence 460 AA;

Query Match 11.0%; Score 13; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 194 GGVLAALAAAYCLS 206
|||||

RESULT 112
AAP90141
ID AAP90141 standard; protein; 460 AA.
XX
XX AAP90141;
AC
XX 25-MAR-2003 (revised)
DT 13-NOV-1989 (first entry)
XX
XX Protein sequence of combined open reading frame of hepatitis c virus cDNA
DE in clones 35, 36, 81 and 32.
XX
XX Hepatitis C virus; combined open reading frame; clone 35; clone 36;
XX clone 81; clone 32; vaccine.
XX
XX Pan troglodytes.
OS
XX GB2212511-A.
XX
XX 26-JUL-1989.
XX
XX 18-NOV-1988; 88GB-00027024.
XX
XX 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
XX (CHIR) CHIRON CORP.
PA
XX Houghton M, Choo QL, Kuo G;
XX
XX WPI; 1989-215054/30.
DR N-PSDB; AAN90310.
XX
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
XX Disclosure; Fig 9-1, 9-2; 30pp; English.
XX
XX The sequence is encoded by the combined open reading frame of hepatitis C
CC virus cDNA (see AAN90310). These antigens react with antibodies in
CC patients with non-A non-B hepatitis (NANBH). They can be used to diagnose
CC HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or
CC to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 460 AA;

Query Match 11.0%; Score 13; DB 1; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.0013;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 194 GGVLAALAAAYCLS 206

RESULT 113
AAY01622
ID AAY01622 standard; protein; 504 AA.
XX AC AAY01622;
XX DT 21-JUN-1999 (first entry)
XX DE Protein encoded by the coding strand of the HCV J1 NS3-NS4 domain.
XX KW HCV; J7 isolate; J1 isolate; HCV1; immunoassay; asiatic strain;
XX KW diagnosis; HCV infection; blood screening; immunisation; antiviral.
XX OS Hepatitis C virus.
XX PN US5871903-A.
XX PD 16-FEB-1999.
XX PF 08-MAY-1995; 95US-00436965.
XX PR 15-SEP-1989; 89US-00408045.
XX PR 21-DEC-1989; 89US-00456142.
XX PR 04-JAN-1991; 91US-00637380.
XX PR 02-AUG-1993; 93US-00101280.
XX PR 24-FEB-1994; 94US-00201066.
XX PR 03-NOV-1994; 94US-00334255.
XX PA (NAHE-) NAT INST OF HEALTH JAPAN.
XX PA (CHIR) CHIRON CORP.
XX PI Saito I, Miyamura T;
XX DR WPI; 1999-166619/14.
XX DR N-PSDB; ANX26741.
XX PT Immunoassays for Asiatic strains of hepatitis C virus - for diagnosis of
XX PT infection and screening blood supplies.
XX PS Disclosure; Fig 16A-B; 43pp; English.
XX CC The present sequence is encoded by the consensus sequence of the coding
XX CC strand of a new hepatitis C virus (HCV), J1, NS3-NS4 domain. The J1 and
XX CC J7 (also a new HCV isolate) isolates comprise sequences which are
XX CC distinct from the prototype HCV isolates, HCV1. The specification
XX CC describes immunoassays for HCV based on antigens from Asiatic strains not
XX CC cross-reactive with HCV-1. The assays are used for diagnosis of HCV
XX CC infection and to screen donated blood. The anti-HCV antibodies are also
XX CC useful therapeutically and prophylactically (passive immunisation); in
XX CC screening for antiviral agents; for isolation, purification and
XX CC identification of non-A, non-B hepatitis virus (e.g. by affinity
XX CC chromatography) and to raise anti-idiotypic antibodies (useful for
XX CC treatment or diagnosis and to determine immunogenic regions of the HCV
XX CC antigens)
XX SQ Sequence 504 AA;

Query Match 11.0%; Score 13; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 402 GGVLAALAAAYCLS 414

RESULT 114
AAR66631
ID AAR66631 standard; protein; 507 AA.
XX AC AAR66631;
XX DT 25-MAR-2003 (revised)
XX DT 31-AUG-1995 (first entry)
XX DE HCV J1 NS3-NS4 domain C200 region consensus protein.
XX KW Hepatitis C virus J1 NS3-NS4 domain C200 region;
XX KW anti-HCV vaccine development; non-A non-B virus; diagnostic polypeptides;
XX KW HCV probes.
XX OS Hepatitis C virus.
XX PN US5372928-A.
XX PD 13-DEC-1994.
XX PF 24-FEB-1994; 94US-00201066.
XX PR 15-SEP-1989; 89US-00408045.
XX PR 21-DEC-1989; 89US-00456142.
XX PR 04-JAN-1991; 91US-00637380.
XX PR 02-AUG-1993; 93US-00101280.
XX PA (CHIR) CHIRON CORP.
XX PA (NAHE-) NAT INST OF HEALTH JAPAN.
XX PI Han J, Saito I, Miyamura T, Cha T, Kolberg JA, Houghton M;
XX PI Irvine BD, Weiner AJ;
XX DR WPI; 1995-030306/04.
XX DR N-PSDB; AAQ79774.
XX PT Method of detecting hepatitis C virus polynucleotide - utilises probe
XX PT based on DNA of new HCV isolates J1 and J7.
XX PS Claim 1; Fig 16; 45pp; English.
XX CC AAQ79774 encodes AAR66631 the prod. of the hepatitis C virus (HCV) J1 NS3
XX CC -NS4 domain C200 region consensus sequence. They can be used to provide
XX CC new oligonucleotides and polypeptides for use in diagnostics, recombinant
XX CC protein prodn. and anti-HCV vaccine development. (Updated on 25-MAR-2003
XX CC to correct PF field.)
XX SQ Sequence 507 AA;

Query Match 11.0%; Score 13; DB 2; Length 507;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 398 GGVLAALAAAYCLS 410

RESULT 115
AAR33565
ID AAR33565 standard; protein; 592 AA.
XX AC AAR33565;
XX DT 25-MAR-2003 (revised)
XX DT 01-JUL-1993 (first entry)
XX DE CKS-HCV antigen fusion protein pHCV-62.
XX KW Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
XX KW immunodot assay; Non-A, non-B hepatitis.
XX XX

```

OS Hepatitis C virus.
XX WO9304087-A1.
XX 04-MAR-1993.
XX 21-AUG-1992; 92WO-US007187.
XX 21-AUG-1991; 91US-00748566.
XX (ABBO ) ABBOTT LAB.
XX Desai SM, Casey JM, Rupprecht KR, Devare SG;
XX WPI; 1993-093940/11.
XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.
XX Claim 1; Page 53-55; 206pp; English.
XX A specific antigenic region of the HCV genome is expressed as a chimeric
XX fusion with E.coli CMP-KDO synthetase (CKS) gene. The fusion protein pHCV
XX -62 can be used to detect antibodies and antigen in body fluids from
XX individuals exposed to HCV e.g. in confirmatory, competition or
XX neutralisation assays. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 592 AA;
Query Match 11.0%; Score 13; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 320 GGVLAALAAAYCLS 332

RESULT 116
AAB69023
ID AAB69023 standard; peptide; 592 AA.
XX
AC AAB69023;
XX
DT 17-APR-2001 (first entry)
XX
DE HCV recombinant antigen pHCV-62 amino acid sequence SEQ ID NO:47.
XX
KW Hepatitis C virus; HCV; antigen; detection; antibody.
XX
OS Hepatitis C virus.
XX
XX US6172189-B1.
XX
PD 09-JAN-2001.
XX
XX 02-JUN-1997; 97US-00867611.
XX
XX 24-AUG-1990; 90US-00572822.
XX 07-NOV-1990; 90US-00614069.
XX 21-AUG-1991; 91US-00748561.
XX 21-AUG-1991; 91US-00748565.
XX 19-NOV-1992; 92US-00989843.
XX 10-JAN-1994; 94US-00179896.
XX 01-MAY-1996; 96US-00646757.
XX (ABBO ) ABBOTT LAB.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
XX Gutierrez RA, Lesniowski RR, Stewart JL, Rupprecht KR;

```

```

DR WPI; 2001-122352/13.
XX
XX New recombinant antigens representing distinct antigenic regions of
XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and
XX antigens in body fluids of individuals exposed to HCV.
XX
XX Example 16; Col 173-178; 167pp; English.
XX
XX The present invention describes recombinant Hepatitis C virus (HCV)
XX antigens (I). (I) is useful as a reagent for the detection of antibodies
XX and antigen in body fluids from individuals exposed to HCV. The HCV assay
XX uses reliable and efficient reagents and methods to accurately detect the
XX presence of HCV antibodies in samples obtained from individuals suspected
XX of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
XX AAB69001 to AAB69032 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 592 AA;
Query Match 11.0%; Score 13; DB 4; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 320 GGVLAALAAAYCLS 332

RESULT 117
ABW01888
ID ABW01888 standard; protein; 592 AA.
XX
XX ABW01888;
XX
DT 12-FEB-2004 (first entry)
XX
DE HCV CKS-C100 recombinant antigen, pHCV-62.
XX
KW Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
KW fusion protein.
XX
XX Chimeric - Hepatitis C virus.
XX Chimeric - Escherichia coli.
XX Chimeric - Unidentified.
XX
XX US6593083-B1.
XX
XX 15-JUL-2003.
XX
XX 17-OCT-2000; 2000US-00690359.
XX
XX 24-AUG-1990; 90US-00572822.
XX 07-NOV-1990; 90US-00614069.
XX 21-AUG-1991; 91US-00748561.
XX 21-AUG-1991; 91US-00748565.
XX 19-NOV-1992; 92US-00989843.
XX 10-JAN-1994; 94US-00179896.
XX 01-MAY-1996; 96US-00646757.
XX 02-JUN-1997; 97US-00867611.
XX (DEVA/) DEVARE S G.
XX (DESA/) DESAI S M.
XX (CASE/) CASEY J M.
XX (DAIL/) DAILEY S H.
XX (DAWS/) DAWSON G J.
XX (GUTI/) GUTIERREZ R A.
XX (LESN/) LESNIEWSKI R R.
XX (STEW/) STEWART J L.
XX (RUPP/) RUPPRECHT K R.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

```


PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
DR WPI; 2003-828264/77.
XX
XX Identifying the presence of an antibody in a fluid sample, where the
PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
PT antigen by contacting the fluid sample, which may contain a HCV antibody
PT with at a polypeptide.
XX
PS Example 16; Col 173-176; 168pp; English.
XX
XX The invention relates to a method for identifying the presence of an
CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
CC The method involves providing a fluid sample containing at least one HCV
CC antibody, contacting the fluid sample with at least one polypeptide or
CC recombinant fusion protein for complexing the antibody with the
CC polypeptide or recombinant fusion protein to provide an antibody-
CC polypeptide complex and detecting the complex. The present sequence is
CC pHCV-62 fusion protein which comprises Escherichia coli Cks (CTP: CMP-3-
CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
CC is used in the invention
XX
XX Sequence 592 AA;
SQ
Query Match 11.0%; Score 13; DB 7; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 320 GGVLAALAAAYCLS 332
RESULT 118
AAR33566
ID AAR33566 standard; protein; 594 AA.
XX
AC AAR33566;
XX
DT 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
XX CKS-HCV antigen fusion protein pHCV-63.
DE
XX Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
KW immunodot assay; Non-A, non-B hepatitis.
KW
XX Hepatitis C virus.
OS
XX WO9304087-A1.
PN
XX 04-MAR-1993.
PD
XX 21-AUG-1992; 92WO-US007187.
PF
XX 21-AUG-1991; 91US-00748566.
PR (ABBO) ABBOTT LAB.
XX
PI Desai SM, Casey JM, Rupprecht KR, Devare SG;
FI WPI; 1993-093940/11.
XX
XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
PT hepatitis C virus.
XX
XX Claim 2; Page 55-57; 206pp; English.
PS
XX A specific antigenic region of the HCV genome is expressed as a chimeric
CC fusion with E.coli CMP-KDO synthetase (CKS) gene. The fusion protein pHCV
CC -63 can be used to detect antibodies and antigen in body fluids from

CC individuals exposed to HCV e.g. in confirmatory, competition or
CC neutralisation assays. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 594 AA;
Query Match 11.0%; Score 13; DB 2; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 322 GGVLAALAAAYCLS 334
RESULT 119
AAB69024
ID AAB69024 standard; peptide; 594 AA.
XX
AC AAB69024;
XX
DT 17-APR-2001 (first entry)
XX
XX HCV recombinant antigen pHCV-63 amino acid sequence SEQ ID NO:48.
DE
XX Hepatitis C virus; HCV; antigen; detection; antibody.
KW
XX Hepatitis C virus.
OS
XX US6172189-B1.
PN
XX 09-JAN-2001.
PD
XX 02-JUN-1997; 97US-00867611.
PF
XX 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
XX
XX (ABBO) ABBOTT LAB.
PA
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2001-122352/13.
DR
XX New recombinant antigens representing distinct antigenic regions of
PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
PT antigens in body fluids of individuals exposed to HCV.
XX
XX Claim 2; Col 177-180; 167pp; English.
PS
XX The present invention describes recombinant Hepatitis C virus (HCV)
CC antigens (I). (I) is useful as a reagent for the detection of antibodies
CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
CC uses reliable and efficient reagents and methods to accurately detect the
CC presence of HCV antibodies in samples obtained from individuals suspected
CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
CC AAB69001 to AAB69032 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 594 AA;
Query Match 11.0%; Score 13; DB 4; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30

```
Db      322 GGVLAALAAAYCLS 334
RESULT 120
ABW01889
XX ID ABW01889 standard; protein; 594 AA.
XX AC ABW01889;
XX DT 12-FEB-2004 (first entry)
XX DE HCV CKS-C100 recombinant antigen, pHCV-63.
XX KW Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
XX KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
XX OS fusion protein.
XX OS Chimeric - Hepatitis C virus.
XX OS Chimeric - Escherichia coli.
XX OS Chimeric - Unidentified.
XX PN US6593083-B1.
XX PD 15-JUL-2003.
XX PF 17-OCT-2000; 2000US-00690359.
XX PR 24-AUG-1990; 90US-00572822.
XX PR 07-NOV-1990; 90US-00614069.
XX PR 21-AUG-1991; 91US-00748561.
XX PR 21-AUG-1991; 91US-00748565.
XX PR 21-AUG-1991; 91US-00748566.
XX PR 19-NOV-1992; 92US-00989843.
XX PR 10-JAN-1994; 94US-00179896.
XX PR 01-MAY-1996; 96US-00646757.
XX PR 02-JUN-1997; 97US-00867611.
XX PA (DEVA/) DEVARE S G.
XX PA (DESA/) DESAI S M.
XX PA (CASE/) CASEY J M.
XX PA (DAILEY/) DAILEY S H.
XX PA (DAWS/) DAWSON G J.
XX PA (GUT/) GUTIERREZ R A.
XX PA (LESN/) LESNIEWSKI R R.
XX PA (STEW/) STEWART J L.
XX PA (RUPP/) RUPPRECHT K R.
XX PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
XX PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2003-828264/77.
XX PT Identifying the presence of an antibody in a fluid sample, where the
XX PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
XX PT antigen by contacting the fluid sample, which may contain a HCV antibody
XX PT with at a polypeptide.
XX PS Example 16; Col 177-180; 168pp; English.
XX CC The invention relates to a method for identifying the presence of an
XX CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
XX CC The method involves providing a fluid sample containing at least one HCV
XX CC antibody, contacting the fluid sample with at least one polypeptide or
XX CC recombinant fusion protein for complexing the antibody with the
XX CC polypeptide or recombinant fusion protein to provide an antibody-
XX CC polypeptide complex and detecting the complex. The present sequence is
XX CC pHCV-63 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
XX CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
XX CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
XX CC is used in the invention
XX SQ Sequence 594 AA;
Query Match 11.0%; Score 13; DB 2; Length 597;

Query Match 11.0%; Score 13; DB 7; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 322 GGVLAALAAAYCLS 334
|||||
|

RESULT 121
AAR21571
XX ID AAR21571 standard; protein; 597 AA.
XX AC AAR21571;
XX DT 10-MAR-2003 (revised)
XX DT 09-JUN-1992 (first entry)
XX DE HCV CKS-C100D1 - pHCV-57.
XX KW Hepatitis C virus; antigen; diagnosis; inhibitor; CMP-KDO synthase; CKS;
XX KW HCV CKS-C100D1; NANBHV; C100.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Peptide 1..239
XX FT /label= CKS
XX FT Region 240..247
XX FT /label= linker
XX FT Peptide 248..587
XX FT /label= C100
XX FT /note= "HCV region NS4, amino acids 1569-1574 and 1598-
XX FT 1931"
XX FT Region 588..597
XX FT /label= linker
XX EP472207-A.
XX PN 26-FEB-1992.
XX PD 23-AUG-1991; 91EP-00114161.
XX PF 24-AUG-1990; 90US-00572822.
XX PR 07-NOV-1990; 90US-00614069.
XX PA (ABBO ) ABBOTT LAB.
XX PI Devare SG, Desai SM, Casey JM, Dawson GJ, Lesniewski RR;
XX PI Dailey SH, Gutierrez RA, Stewart JL;
XX WPI; 1992-066430/09.
XX DR N-PSDB; AAQ21684.
XX PT Recombinant hepatitis C virus antigens - produced as fusion proteins and
XX PT representing distinct antigenic regions of the HCV genome.
XX PS Disclosure; Fig 42, Page 85-87; 115pp; English.
XX CC pHCV-57 fusion protein (mol.wt. 65 kD) expresses at significantly higher
XX CC levels than the pHCV-24 fusion protein. pHCV-57 comprises a 23 amino acid
XX CC deletion (HCV amino acids 1575-1597) in the extreme N-terminal portion of
XX CC the HCV C100 region, compared to the pHCV-24 fusion protein. The
XX CC polypeptide represents a distinct antigenic region of the HCV genome and
XX CC can be used for the detection of antibodies and antigens for early
XX CC diagnosis of HCV infection. The polypeptide can also be used to develop
XX CC specific inhibitors of viral replication and for therapeutic purposes.
XX CC (Updated on 10-MAR-2003 to add missing OS field.)
XX SQ Sequence 597 AA;
```

```
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 320 GGVLAALAAAYCLS 332

RESULT 122
AAR33638
ID AAR33638 standard; protein; 597 AA.
AC AAR33638;
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 13-JUL-1993 (first entry)
XX
XX HCV C100D1 recombinant antigen encoded by pHCV-57.
XX
XX Hepatitis C virus; NANBH; non-A, non-B hepatitis virus; NANBH;
KW non-structural protein; pHCV-57; diagnosis; CKS fusion protein;
KW CTP:CMF-3-deoxy-manno-octulosonate cytidyl transferase; immunoassay.
XX
XX Hepatitis C virus; Virus.
XX
XX WO9304089-A1.
XX
XX 04-MAR-1993.
XX
XX 21-AUG-1992; 92WO-US006964.
XX
XX 21-AUG-1991; 91US-00748565.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Desai SM, Dailey SH, Devare SG;
XX
XX WPI; 1993-093942/11.
XX
XX N-PSDB; AAQ38272.
XX
XX New recombinant NS5 region antigens - for hepatitis C assay for detecting
PT hepatitis C virus infections.
XX
XX Example 8; Page 78-80; 164pp; English.
XX
XX The clone pHCV-57 was constructed to overcome poor expression levels of
CC the HCV CKS-c100 recombinant antigen. HCV CKS-c100 antigen consists of
CC 239 amino acids of CKS, eight amino acids contributed by linker DNA
CC sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1931)
CC and 10 additional amino acids contributed by linker DNA sequences. A 69bp
CC Ddel fragment was deleted from pHCV-24 (a plasmid which expresses the HCV
CC CKS-c100 antigen under control of the lac promoter). The pHCV-57 fusion
CC protein (containing a 23 amino acid deletion, i.e. of HCV amino acids
CC 1575-1597) was expressed at a significantly higher level than the pHCV-24
CC fusion protein. (Updated on 25-MAR-2003 to correct PN field.) (Updated on
CC 24-OCT-2003 to standardise OS field)
XX
XX Sequence 597 AA;
XX
XX Query Match 11.0%; Score 13; DB 2; Length 597;
XX Best Local Similarity 100.0%; Pred. No. 0.0016;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 320 GGVLAALAAAYCLS 332

RESULT 123
AAR33600
ID AAR33600 standard; protein; 597 AA.
XX
```

```
AAR33600;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-JUL-1993 (first entry)
XX
XX HCV C100D1 recombinant antigen encoded by pHCV-57.
XX
XX Hepatitis C virus; non-A, non-B hepatitis virus; NANBH;
KW non-structural protein; CMF-KDO synthetase; CKS fusion protein;
KW CTP:CMF-3-deoxy-manno-octulosonate cytidyl transferase; immunoassay;
XX pHCV-57.
XX
XX Hepatitis C virus; Virus.
XX
XX WO9304088-A1.
XX
XX 04-MAR-1993.
XX
XX 21-AUG-1992; 92WO-US007188.
XX
XX 21-AUG-1991; 91US-00748561.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Dailey SH, Desai SM, Devare SG;
XX
XX WPI; 1993-093941/11.
XX
XX N-PSDB; AAQ38257.
XX
XX Hepatitis C assay using recombinant NS1 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.
XX
XX Example 8; Page 88-90; 175pp; English.
XX
XX The clone pHCV-57 was constructed to overcome poor expression levels of
CC the HCV CKS-c100 recombinant antigen. HCV CKS-c100 antigen consists of
CC 239 amino acids of CKS, eight amino acids contributed by linker DNA
CC sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1931)
CC and 10 additional amino acids contributed by linker DNA sequences. A 69bp
CC Ddel fragment was deleted from pHCV-24 (a plasmid which expresses the HCV
CC CKS-c100 antigen under control of the lac promoter). The pHCV-57 fusion
CC protein (containing a 23 amino acid deletion, i.e. of HCV amino acids
CC 1575-1597) was expressed at a significantly higher level than the pHCV-24
CC fusion protein. (Updated on 25-MAR-2003 to correct PN field.) (Updated on
CC 24-OCT-2003 to standardise OS field)
XX
XX Sequence 597 AA;
XX
XX Query Match 11.0%; Score 13; DB 2; Length 597;
XX Best Local Similarity 100.0%; Pred. No. 0.0016;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 320 GGVLAALAAAYCLS 332

RESULT 124
AAR33580
ID AAR33580 standard; protein; 597 AA.
XX
XX AAR33580;
XX
XX 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
XX HCV-C100D1 recombinant antigen encoded by pHCV-57.
XX
XX Hepatitis C virus; NS3; C100 antigen; CKS fusion protein;
KW CMF-KDO synthetase; immunodot assay; Non-A, non-B hepatitis.
XX
```

```

OS Hepatitis C virus.
XX WO9304087-A1.
XX 04-MAR-1993.
XX
XX 21-AUG-1992; 92WO-US007187.
XX
XX 21-AUG-1991; 91US-00748566.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Desai SM, Casey JM, Rupprecht KR, Devare SG;
XX WPI; 1993-093940/11.
XX N-PSDB; AAF38242.
XX
XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.
XX
XX Example 8; Page 114-116; 206pp; English.
XX
XX The clone pHCV-57 was constructed to overcome poor expression levels of
XX the HCV CKS-ci00 recombinant antigen. HCV CKS-ci00 antigen consists of
XX 239 amino acids of CKS, eight amino acids contributed by linker DNA
XX sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1931)
XX and 10 additional amino acids contributed by linker DNA sequences. A 69bp
XX DdeI fragment was deleted from pHCV-24 (a plasmid which expresses the HCV
XX CKS-ci00 antigen under control of the lac promoter). The pHCV-57 fusion
XX protein (containing a 23 amino acid deletion, i.e. of HCV amino acids
XX 1575-1597) was expressed at a significantly higher level than the pHCV-24
XX fusion protein. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 597 AA;
XX
XX Query Match 11.0%; Score 13; DB 2; Length 597;
XX Best Local Similarity 100.0%; Pred. No. 0.0016;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAAALAAAYCLS 30
XX |||||
XX DB 320 GGVLAAALAAAYCLS 332
XX
XX RESULT 125
XX AAB51378
XX ID AAB51378 standard; protein; 597 AA.
XX AC AAB51378;
XX
XX 17-APR-2001 (first entry)
XX
XX HCV recombinant antigen pHCV-57 protein sequence SEQ ID NO:16.
XX
XX Hepatitis C virus; HCV; antigen; detection; antibody.
XX
XX Hepatitis C virus.
XX
XX US6172189-B1.
XX
XX 09-JAN-2001.
XX
XX 02-JUN-1997; 97US-00867611.
XX
XX 24-AUG-1990; 90US-00572822.
XX 07-NOV-1990; 90US-00614069.
XX 21-AUG-1991; 91US-00748561.
XX 21-AUG-1991; 91US-00748565.
XX 21-AUG-1991; 91US-00748566.
XX 19-NOV-1992; 92US-00989843.
XX 10-JAN-1994; 94US-00179896.
XX 01-MAY-1996; 96US-00646757.
XX 01-MAY-1996; 96US-00646757.
XX
XX (DEVA/) DEVARE S G.
XX (DESA/) DESAI S M.
XX (CASE/) CASEY J M.
XX (DAIL/) DAILEY S H.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2001-122352/13.
XX N-PSDB; AAF32225.
XX
XX New recombinant antigens representing distinct antigenic regions of
XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and
XX antigens in body fluids of individuals exposed to HCV.
XX
XX Example 8; Col 91-94; 167pp; English.
XX
XX The present invention describes recombinant Hepatitis C virus (HCV)
XX antigens (I). (I) is useful as a reagent for the detection of antibodies
XX and antigen in body fluids from individuals exposed to HCV. The HCV assay
XX uses reliable and efficient reagents and methods to accurately detect the
XX presence of HCV antibodies in samples obtained from individuals suspected
XX of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
XX AAB69001 to AAB69032 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 597 AA;
XX
XX Query Match 11.0%; Score 13; DB 4; Length 597;
XX Best Local Similarity 100.0%; Pred. No. 0.0016;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAAALAAAYCLS 30
XX |||||
XX DB 320 GGVLAAALAAAYCLS 332
XX
XX RESULT 126
XX ABW01864
XX ID ABW01864 standard; protein; 597 AA.
XX AC ABW01864;
XX
XX 12-FEB-2004 (first entry)
XX
XX HCV-CKS-C100D1 recombinant antigen, pHCV-57.
XX
XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
XX CTP: CMP-3-deoxy-manno-octulose cytidyl transferase; antigen;
XX fusion protein.
XX
XX Chimeric - Hepatitis C virus.
XX OS Chimeric - Escherichia coli.
XX OS Chimeric - Unidentified.
XX
XX US6593083-B1.
XX
XX 15-JUL-2003.
XX
XX 17-OCT-2000; 2000US-00690359.
XX
XX 24-AUG-1990; 90US-00572822.
XX 07-NOV-1990; 90US-00614069.
XX 21-AUG-1991; 91US-00748561.
XX 21-AUG-1991; 91US-00748565.
XX 21-AUG-1991; 91US-00748566.
XX 19-NOV-1992; 92US-00989843.
XX 10-JAN-1994; 94US-00179896.
XX 01-MAY-1996; 96US-00646757.
XX 02-JUN-1997; 97US-00867611.
XX
XX (DEVA/) DEVARE S G.
XX (DESA/) DESAI S M.
XX (CASE/) CASEY J M.
XX (DAIL/) DAILEY S H.

```

PA (DAWS/) DAWSON G J.
 PA (GUTI/) GUTIERREZ R A.
 PA (LESN/) LESNIEWSKI R R.
 PA (STEW/) STEWART J L.
 PA (RUPP/) RUPPRECHT K R.
 XX
 XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 XX N-PSDB; AAD63437.
 DR WPI; 2003-828264/77.
 DR N-PSDB; AAD63437.
 XX
 XX Identifying the presence of an antibody in a fluid sample, where the
 PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
 PT antigen by contacting the fluid sample, which may contain a HCV antibody
 PT with at a polypeptide.
 XX
 XX Example 8; Col 91-94; 168pp; English.
 XX
 XX The invention relates to a method for identifying the presence of an
 CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
 CC The method involves providing a fluid sample containing at least one HCV
 CC antibody, contacting the fluid sample with at least one polypeptide or
 CC recombinant fusion protein for complexing the antibody with the
 CC polypeptide or recombinant fusion protein to provide an antibody-
 CC polypeptide complex and detecting the complex. The present sequence is
 CC pHCV-57 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
 CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
 CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
 CC is used in the invention
 XX
 XX Sequence 597 AA;
 SQ
 Query Match 11.0%; Score 13; DB 7; Length 597;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 18 GGVLAAALAAAYCLIS 30
 |||||
 Db 320 GGVLAAALAAAYCLIS 332
 |||||
 RESULT 127
 AAR21572
 ID AAR21572 standard; protein; 599 AA.
 XX
 AC AAR21572;
 XX
 DT 10-MAR-2003 (revised)
 DT 09-JUN-1992 (first entry)
 XX
 DE HCV CKS-C100D2 - pHCV-58.
 XX
 KW Hepatitis C virus; antigen; diagnosis; inhibitor; CMP-KDO synthase; CKS;
 KW HCV CKS-C100D2; NANBH; C100.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..239
 FT /label= CKS
 FT Region 240..247
 FT /label= linker
 FT Peptide 248..589
 FT /label= C100
 FT /note= "HCV region NS4, amino acids 1569-1599 and 1621-
 FT 1931"
 FT Region 590..599
 FT /label= linker
 XX
 XX EP472207-A.
 PN
 XX

PD 26-FEB-1992.
 XX
 XX 23-AUG-1991; 91EP-00114161.
 XX
 XX 24-AUG-1990; 90US-00572822.
 PR 07-NOV-1990; 90US-00614069.
 XX
 XX (ABBO) ABBOTT LAB.
 XX
 XX Devare SG, Desai SM, Casey JM, Dawson GJ, Lesniewski RR;
 PI Dailey SH, Gutierrez RA, Stewart JL;
 XX WPI; 1992-066430/09.
 DR N-PSDB; AAQ21685.
 XX
 XX Recombinant hepatitis C virus antigens - produced as fusion proteins and
 PT representing distinct antigenic regions of the HCV genome.
 XX
 XX Disclosure; Fig 43, Page 88-90; 115pp; English.
 XX
 XX pHCV-58 fusion protein (mol.wt. 66 kD) expresses at significantly higher
 CC levels than the pHCV-24 fusion protein. pHCV-58 comprises a 21 amino acid
 CC deletion (HCV amino acids 1600-1620) in the extreme N-terminal portion of
 CC the HCV C100 region, compared to the pHCV-24 fusion protein. The
 CC polypeptide represents a distinct antigenic region of the HCV genome and
 CC can be used for the detection of antibodies and antigens for early
 CC diagnosis of HCV infection. The polypeptide can also be used to develop
 CC specific inhibitors of viral replication and for therapeutic purposes.
 CC (Updated on 10-MAR-2003 to add missing OS field.)
 XX
 XX Sequence 599 AA;
 SQ
 Query Match 11.0%; Score 13; DB 2; Length 599;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 18 GGVLAAALAAAYCLIS 30
 |||||
 Db 322 GGVLAAALAAAYCLIS 334
 |||||
 RESULT 128
 AAR33639
 ID AAR33639 standard; protein; 599 AA.
 XX
 AC AAR33639;
 XX
 DT 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 13-JUL-1993 (first entry)
 XX
 DE HCV C100D2 recombinant antigen encoded by pHCV-58.
 XX
 KW Hepatitis C virus; NANBH; non-A, non-B hepatitis; CMP-KDO synthetase;
 KW non-structural protein; pHCV-58; diagnosis; CKS fusion protein;
 KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; immunoassay.
 XX
 OS Hepatitis C virus; Virus.
 OS
 XX WO9304089-A1.
 PN
 XX 04-MAR-1993.
 PD
 XX 21-AUG-1992; 92WO-US006964.
 PF
 XX 21-AUG-1991; 91US-00748565.
 PR
 XX (ABBO) ABBOTT LAB.
 PA
 XX Desai SM, Dailey SH, Devare SG;
 PI
 XX WPI; 1993-093942/11.
 DR N-PSDB; AAQ38273.
 DR

XX New recombinant NS5 region antigens - for hepatitis C assay for detecting
PT hepatitis C virus infections.

XX Example 8; Page 83-85; 164pp; English.

XX The clone pHCV-58 was constructed to overcome poor expression levels of
CC the HCV CKS-c100 recombinant antigen. HCV CKS-c100 antigen consists of
CC 239 amino acids of CKS, eight amino acids contributed by linker DNA
CC sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1931)
CC and 10 additional amino acids contributed by linker DNA sequences. A 63bp
CC NlaIV-HaeIII fragment was deleted from pHCV-24 (a plasmid which expresses
CC the HCV CKS- c100 antigen under control of the lac promoter). The pHCV-58
CC fusion protein (containing a 21 amino acid deletion, i.e. of HCV amino
CC acids 1600-1620) was expressed at a significantly higher level than the
CC pHCV-24 fusion protein. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 599 AA;

Query Match 11.0%; Score 13; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
Db 322 GGVLAAALAAAYCLS 334

RESULT 129

AAR33601
ID AAR33601 standard; protein; 599 AA.

AC AAR33601;

DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-JUL-1993 (first entry)

XX HCV C100D2 recombinant antigen encoded by pHCV-58.

XX Hepatitis C virus; non-A, non-B hepatitis virus; NANBH;

XX non-structural protein; CMP-KDO synthetase; CKS fusion protein;

KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; immunoassay;

KW pHCV-58.

XX Hepatitis C virus; Virus.

OS WO9304088-A1.

PN 04-MAR-1993.

XX 21-AUG-1992; 92WO-US007188.

XX 21-AUG-1991; 91US-00748561.

XX (ABBO) ABBOTT LAB.

PI Dailey SH, Desai SM, Devare SG;

XX WPI; 1993-093941/11.

DR N-PSDB; AAQ38258.

XX Hepatitis C assay using recombinant NS1 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
PT hepatitis C virus.
XX Example 8; Page 93-95; 175pp; English.

XX The clone pHCV-58 was constructed to overcome poor expression levels of
CC the HCV CKS-c100 recombinant antigen. HCV CKS-c100 antigen consists of
CC 239 amino acids of CKS, eight amino acids contributed by linker DNA
CC sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1931)

CC and 10 additional amino acids contributed by linker DNA sequences. A 63bp
CC NlaIV-HaeIII fragment was deleted from pHCV-24 (a plasmid which expresses
CC the HCV CKS- c100 antigen under control of the lac promoter). The pHCV-58
CC fusion protein (containing a 21 amino acid deletion, i.e. of HCV amino
CC acids 1600-1620) was expressed at a significantly higher level than the
CC pHCV-24 fusion protein. (Updated on 25-MAR-2003 to correct PN field.)
CC (Updated on 24-OCT-2003 to standardise OS field)

XX Sequence 599 AA;

Query Match 11.0%; Score 13; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
Db 322 GGVLAAALAAAYCLS 334

RESULT 130

AAR33581

ID AAR33581 standard; protein; 599 AA.

AC AAR33581;

DT 25-MAR-2003 (revised)

DT 01-JUL-1993 (first entry)

XX HCV C100D2 recombinant antigen encoded by pHCV-58.

XX Hepatitis C virus; NS3; C100 antigen; CKS fusion protein;

KW CMP-KDO synthetase; immunodot assay; Non-A, non-B hepatitis.

XX Hepatitis C virus.

XX WO9304087-A1.

XX 04-MAR-1993.

XX 21-AUG-1992; 92WO-US007187.

XX 21-AUG-1991; 91US-00748566.

XX (ABBO) ABBOTT LAB.

XX Desai SM, Casey JM, Rupprecht KR, Devare SG;

XX WPI; 1993-093940/11.

XX N-PSDB; AAQ38243.

XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
PT hepatitis C virus.
XX Example 8; Page 119-121; 206pp; English.

XX The clone pHCV-58 was constructed to overcome poor expression levels of
CC the HCV CKS-c100 recombinant antigen. HCV CKS-c100 antigen consists of
CC 239 amino acids of CKS, eight amino acids contributed by linker DNA
CC sequences, 363 amino acids from HCV NS4 region (amino acids 1569-1931)
CC and 10 additional amino acids contributed by linker DNA sequences. A 63bp
CC NlaIV-HaeIII fragment was deleted from pHCV-24 (a plasmid which expresses
CC the HCV CKS- c100 antigen under control of the lac promoter). The pHCV-58
CC fusion protein (containing a 21 amino acid deletion, i.e. of HCV amino
CC acids 1600-1620) was expressed at a significantly higher level than the
CC pHCV-24 fusion protein. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 599 AA;

Query Match 11.0%; Score 13; DB 2; Length 599;

Best Local Similarity 100.0%; Pred. No. 0.0016;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 322 GGVLAALAAAYCLS 334

RESULT 131

ABW01865
ID AAB51379 standard; protein; 599 AA.

XX AC AAB51379;

XX DT 17-APR-2001 (first entry)

XX DE HCV recombinant antigen pHCV-58 protein sequence SEQ ID NO:18.

XX KW Hepatitis C virus; HCV; antigen; detection; antibody.

XX OS Hepatitis C virus.

XX PN US6172189-B1.

XX PD 09-JAN-2001.

XX PF 02-JUN-1997; 97US-00867611.

XX PR 24-AUG-1990; 90US-00572822.

XX PR 07-NOV-1990; 90US-00614069.

XX PR 21-AUG-1991; 91US-00748561.

XX PR 21-AUG-1991; 91US-00748565.

XX PR 21-AUG-1991; 91US-00748566.

XX PR 19-NOV-1992; 92US-00989843.

XX PR 10-JAN-1994; 94US-00179896.

XX PR 01-MAY-1996; 96US-00646757.

XX PA (ABBO) ABBOTT LAB.

XX PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX DR WPI; 2001-122352/13.

XX DR N-PSDB; AAF32226.

XX PT New recombinant antigens representing distinct antigenic regions of
PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
PT antigens in body fluids of individuals exposed to HCV.
XX Claim 2; Col 97-102; 167pp; English.
XX The present invention describes recombinant Hepatitis C virus (HCV)
CC antigens (I). (I) is useful as a reagent for the detection of antibodies
CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
CC uses reliable and efficient reagents and methods to accurately detect the
CC presence of HCV antibodies in samples obtained from individuals suspected
CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
CC AAB69001 to AAB69032 represent sequences used in the exemplification of
CC the present invention
XX Sequence 599 AA;

SQ

Query Match 11.0%; Score 13; DB 4; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

DB 322 GGVLAALAAAYCLS 334

RESULT 132

ABW01865
ID ABW01865 standard; protein; 599 AA.

XX AC ABW01865;

XX DT 12-FEB-2004 (first entry)

XX DE HCV-CKS-C100D2 recombinant antigen, pHCV-58.

XX KW Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
fusion protein.

XX OS Chimeric - Hepatitis C virus.

XX OS Chimeric - Escherichia coli.

XX OS Chimeric - Unidentified.

XX PN US6593083-B1.

XX PD 15-JUL-2003.

XX PF 17-OCT-2000; 2000US-00690359.

XX PR 24-AUG-1990; 90US-00572822.

XX PR 07-NOV-1990; 90US-00614069.

XX PR 21-AUG-1991; 91US-00748561.

XX PR 21-AUG-1991; 91US-00748565.

XX PR 21-AUG-1991; 91US-00748566.

XX PR 19-NOV-1992; 92US-00989843.

XX PR 10-JAN-1994; 94US-00179896.

XX PR 01-MAY-1996; 96US-00646757.

XX PR 02-JUN-1997; 97US-00867611.

XX PA (DEVA/) DEVARE S G.

XX PA (DESA/) DESAI S M.

XX PA (CASE/) CASEY J M.

XX PA (DAIL/) DAILEY S H.

XX PA (DAWS/) DAWSON G J.

XX PA (GUTI/) GUTIERREZ R A.

XX PA (LESN/) LESNIEWSKI R R.

XX PA (STEW/) STEWART J L.

XX PA (RUPP/) RUPPRECHT K R.

XX PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX DR WPI; 2003-828264/77.

XX DR N-PSDB; AAD63438.

XX PT Identifying the presence of an antibody in a fluid sample, where the
PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
PT antigen by contacting the fluid sample, which may contain a HCV antibody
with at a polypeptide.
XX Claim 1; Col 97-102; 168pp; English.

XX The invention relates to a method for identifying the presence of an
CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
CC The method involves providing a fluid sample containing at least one HCV
CC antibody, contacting the fluid sample with at least one polypeptide or
CC recombinant fusion protein for complexing the antibody with the
CC polypeptide or recombinant fusion protein to provide an antibody-
CC polypeptide complex and detecting the complex. The present sequence is
CC pHCV-58 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
CC is used in the invention
XX Sequence 599 AA;

QY 18 GGVLAALAAAYCLS 30

DB 322 GGVLAALAAAYCLS 334

Query Match 11.0%; Score 13; DB 7; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 133
AAR33567
ID AAR33567 standard; protein; 613 AA.
XX
XX
AC AAR33567;
XX
XX 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
XX CKS-HCV antigen fusion protein pHCV-204.
XX
XX Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
KW immunodot assay; Non-A, non-B hepatitis.
XX
XX Hepatitis C virus.
XX
XX WO9304087-A1.
XX
XX 04-MAR-1993.
XX
XX 21-AUG-1992; 92WO-US007187.
XX
XX 21-AUG-1991; 91US-00748566.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Desai SM, Casey JM, Rupprecht KR, Devare SG;
PI WPI; 1993-093940/11.
XX
XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX antibodies and antigen in body fluids from individuals exposed to
XX hepatitis C virus.
XX
XX Claim 1; Page 57-59; 206pp; English.
XX
XX A specific antigenic region of the HCV genome is expressed as a chimeric
XX fusion with E.coli CMP-KDO synthetase (CKS) gene. The fusion protein pHCV
XX -204 can be used to detect antibodies and antigen in body fluids from
XX individuals exposed to HCV e.g. in confirmatory, competition or
XX neutralisation assays. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 613 AA;
SQ
    Query Match      11.0%; Score 13; DB 2; Length 613;
    Best Local Similarity 100.0%; Pred. No. 0.0016;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 341 GGVLAALAAAYCLS 353

RESULT 134
AAB69025
ID AAB69025 standard; peptide; 613 AA.
XX
XX AAB69025;
XX
XX 17-APR-2001 (first entry)
DT
XX
XX HCV recombinant antigen pHCV-204 amino acid sequence SEQ ID NO:49.
DE
XX
XX Hepatitis C virus; HCV; antigen; detection; antibody.
XX
XX Hepatitis C virus.
OS
XX
XX US6172189-B1.
XX
XX 09-JAN-2001.
PD
XX

```

```

PF 02-JUN-1997; 97US-00867611.
XX
XX 24-AUG-1990; 90US-00572822.
PR
XX 07-NOV-1990; 90US-00614069.
PR
XX 21-AUG-1991; 91US-00748561.
PR
XX 21-AUG-1991; 91US-00748565.
PR
XX 21-AUG-1991; 91US-00748566.
PR
XX 19-NOV-1992; 92US-00989843.
PR
XX 10-JAN-1994; 94US-00179896.
PR
XX 01-MAY-1996; 96US-00646757.
XX
XX (ABBO ) ABBOTT LAB.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI Gutierrez RA, Lesniowski RR, Stewart JL, Rupprecht KR;
XX WPI; 2001-122352/13.
XX
XX New recombinant antigens representing distinct antigenic regions of
XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and
XX antigens in body fluids of individuals exposed to HCV.
XX
XX Claim 1; Col 179-184; 167pp; English.
XX
XX The present invention describes recombinant Hepatitis C virus (HCV)
XX antigens (I). (I) is useful as a reagent for the detection of antibodies
XX and antigen in body fluids from individuals exposed to HCV. The HCV assay
XX uses reliable and efficient reagents and methods to accurately detect the
XX presence of HCV antibodies in samples obtained from individuals suspected
XX of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
XX AAB69001 to AAB69032 represent sequences used in the exemplification of
XX the present invention
XX
XX Sequence 613 AA;
SQ
    Query Match      11.0%; Score 13; DB 4; Length 613;
    Best Local Similarity 100.0%; Pred. No. 0.0016;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 341 GGVLAALAAAYCLS 353

RESULT 135
ABW01890
ID ABW01890 standard; protein; 613 AA.
XX
XX AC ABW01890;
XX
XX 12-FEB-2004 (first entry)
DT
XX
XX HCV CKS-C100 recombinant antigen, pHCV-204.
DE
XX
XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
XX fusion protein.
XX
XX Chimeric - Hepatitis C virus.
OS
XX Chimeric - Escherichia coli.
OS
XX Chimeric - Unidentified.
XX
XX US6593083-B1.
PN
XX
XX 15-JUL-2003.
PD
XX
XX 17-OCT-2000; 2000US-00690359.
PF
XX
XX 24-AUG-1990; 90US-00572822.
PR
XX 07-NOV-1990; 90US-00614069.
PR
XX 21-AUG-1991; 91US-00748561.
PR
XX 21-AUG-1991; 91US-00748565.
PR
XX 21-AUG-1991; 91US-00748566.
PR

```


PR 19-NOV-1992; 92US-00989843.
 PR 10-JAN-1994; 94US-00179896.
 PR 01-MAY-1996; 96US-00646757.
 PR 02-JUN-1997; 97US-00867611.
 XX
 PA (DEVA/) DEVARE S G.
 PA (DESA/) DESAI S M.
 PA (CASE/) CASEY J M.
 PA (DAIL/) DAILY S H.
 PA (DAWS/) DAWSON G J.
 PA (GUTI/) GUTIERREZ R A.
 PA (LESN/) LESNIEWSKI R R.
 PA (STEW/) STEWART J L.
 PA (RUPP/) RUPPRECHT K R.
 XX
 XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 XX
 XX WPI; 2003-828264/77.
 XX
 PT Identifying the presence of an antibody in a fluid sample, where the
 PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
 PT antigen by contacting the fluid sample, which may contain a HCV antibody
 PT with at a polypeptide.
 XX
 XX Example 16; Col 179-184; 169pp; English.
 PS
 CC The invention relates to a method for identifying the presence of an
 CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
 CC The method involves providing a fluid sample containing at least one HCV
 CC antibody, contacting the fluid sample with at least one polypeptide or
 CC recombinant fusion protein for complexing the antibody with the
 CC polypeptide or recombinant fusion protein to provide an antibody-
 CC polypeptide complex and detecting the complex. The present sequence is
 CC pHCV-204 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
 CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
 CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
 CC is used in the invention
 XX
 XX Sequence 613 AA;
 SQ
 Query Match 11.0%; Score 13; DB 7; Length 613;
 Best Local Similarity 100.0%; Pred. No. 0.0016;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCIS 30
 DB |||||
 341 GGVLAALAAAYCIS 353
 RESULT 136
 AAW71271
 ID AAW71271 standard; protein; 685 AA.
 XX
 AC AAW71271;
 XX
 DT 17-NOV-1998 (first entry)
 XX
 DE Protein encoded by the partial DNA sequence of a HCV genome, strain H.
 XX
 KW HCV NS3 protease; NS4A cofactor protein; HCV therapy; fusion protein.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9837180-A2.
 XX
 XX 27-AUG-1998.
 PD
 PF 20-FEB-1998; 98WO-US003367.
 XX
 XX 22-FEB-1997; 97US-00804266.
 XX
 XX (ABBO) ABBOTT LAB.

XX Chen C, Molla A, Tripathi RL;
 XX
 DR WPI; 1998-467551/40.
 DR N-PSDB; AAV54955.
 XX
 FT New hepatitis C virus fusion proteins - comprises NS3 protease and NS4A
 PT co-factor, used in assays for screening for compounds for use in HCV
 PT therapy.
 XX
 PS Disclosure; Fig 1a-c; 31pp; English.
 XX
 CC The present sequence is encoded by the partial nucleotide sequence of a
 CC Hepatitis C virus (HCV) genome, strain H. The specification describes a
 CC construct which encodes HCV NS3 protease and NS4A cofactor protein. A non
 CC -autocleavable fusion protein of HCV NS3 protease and HCV NS4A cofactor
 CC protein is produced upon expression, which is biologically active. The
 CC product can be used to obtain drugs which can inhibit NS3 protease
 CC activity for use in HCV therapy. They can also be used to design
 CC compounds which interact with or inhibit the NS3/NS4A fusion proteins
 SQ Sequence 685 AA;
 XX
 Query Match 11.0%; Score 13; DB 2; Length 685;
 Best Local Similarity 100.0%; Pred. No. 0.0018;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCIS 30
 DB |||||
 638 GGVLAALAAAYCIS 650
 RESULT 137
 AAB62633
 ID AAB62633 standard; protein; 686 AA.
 XX
 AC AAB62633;
 XX
 DT 23-JUL-2001 (first entry)
 XX
 DE HCV NS34A polypeptide.
 XX
 KW HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
 KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
 KW immunotherapy; NS34A.
 XX
 OS Synthetic.
 OS Hepatitis C virus.
 XX
 PN WO200138360-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 22-NOV-2000; 2000WO-US032326.
 XX
 PR 24-NOV-1999; 99US-0167502P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Coit D, Medina-Selby A, Selby M, Houghton M;
 XX
 DR WPI; 2001-343948/36.
 DR N-PSDB; AAF83669.
 XX
 XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
 PT a vaccine against HCV, comprises a polypeptide having a mutation that
 PT functionally disrupts the catalytic domain of NS3.
 XX
 PS Disclosure; Fig 9; 340pp; English.
 XX
 CC The invention relates to an isolated mutant non-structural (NS) Hepatitis
 CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
 CC the catalytic domain of NS3, where the mutation functionally disrupts the

CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC in vaccines against HCV and as diagnostics. The antibodies raised against
CC these polypeptides can also be used as diagnostics, or for passive
CC immunotherapy. The antibodies are also useful for isolating and
CC identifying HCV particles. The present sequence represents the amino acid
CC sequence of a NS3/4A polypeptide from the plasmid pCMV-NS3/4A
XX
XX
SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 4; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
|||||

RESULT 138
AAU76377
ID AAU76377 standard; protein; 686 AA.
XX
AC AAU76377;
XX

DT 08-MAY-2002 (first entry)
XX

XX Hepatitis C virus NS3/4a conformational epitope protein sequence.

XX Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion;
KW immunoassay solid support; multiple epitope fusion antigen; MEFA;
KW non-structural protein; mutant; mutein.
XX

OS Hepatitis C virus.
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 403 /note= "Wild-type Thr substituted by Pro"

FT Misc-difference 404 /note= "Wild-type Ser substituted by Ile"

XX WO200196870-A2.

XX 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US019156.

XX 15-JUN-2000; 2000US-0212082P.

XX 02-APR-2001; 2001US-0280811P.

XX 02-APR-2001; 2001US-0280867P.

XX (CHIR) CHIRON CORP.

XX Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;

XX WPI; 2002-090228/12.

XX N-PSDB; ABK15344.

XX Immunoassay solid support, useful for detecting hepatitis C virus
PT infection in biological sample, comprises HCV NS3/4a conformational
PT epitope and multiple epitope fusion antigen bound to the support.

XX Claim 5; Fig 3; 92pp; English.

XX The present invention relates to a new immunoassay solid support
CC consisting essentially of at least one hepatitis C virus (HCV) NS3/4a
CC conformational epitope and a multiple epitope fusion antigen (MEFA),
CC bound to the support. The NS3/4a conformational epitope and/or MEFA,
CC reacts specifically with anti-HCV antibodies present in a biological
CC sample from an HCV-infected individual. The immunoassay of the invention

CC is useful for detecting hepatitis C virus infection in a biological
CC sample. The method of the invention provides a sensitive, accurate
CC diagnostic and prognostic tool to provide adequate patient care and to
CC prevent transmission of HCV by blood and by blood products, or by
CC personal contact. Use of NS3/4a conformational epitope in combination
CC with MEFA, provides a sensitive and reliable method for detecting early
CC HCV seroconversion. Use of MEFA has the added advantages of decreasing
CC masking problems, improving sensitivity in detecting antibodies by
CC allowing a greater number of epitopes on a unit surface area of
CC substrate, and improving substrate. Detection accuracy is increased and
CC the incidence of false results is reduced because of the identification
CC and the use of highly immunogenic HCV antigens which are present during
CC the early stages of HCV seroconversion. The present amino acid sequence
CC represents the non-structural protein NS3/4a conformational epitope of
XX the invention

SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
|||||

RESULT 139

AAE18689

ID AAE18689 standard; protein; 686 AA.

XX
AC AAE18689;

XX
DT 17-MAY-2002 (first entry)

XX HCV-1 NS3/4a mutant conformational antigen.

XX Hepatitis C virus; NS3/4a antigen; HCV infection; mutant; mutein.

XX Hepatitis C virus type 1.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 403 /note= "Wild type Thr substituted with Pro"

FT Misc-difference 404 /note= "Wild type Ser substituted with Ile"

XX WO200196875-A2.

XX 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US019369.

XX 15-JUN-2000; 2000US-0212082P.

XX 02-APR-2001; 2001US-0280811P.

XX 02-APR-2001; 2001US-0280867P.

XX (CHIR) CHIRON CORP.

XX Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;

XX WPI; 2002-179522/23.

XX N-PSDB; AAD29795.

XX Immunoassay solid support useful for detecting hepatitis C virus
PT infection in a biological sample, comprises at least one of HCV anti-core
PT antibody and HCV NS3/4a epitope, bound to the support.

XX Example 2; Fig 4; 87pp; English.

XX The present invention relates to hepatitis C virus (HCV) core antigen and

CC NS (nonstructural) 3/4a antibody combination assay that can detect both
CC HCV antigens and antibodies present in a sample using a single solid
CC matrix as well as immunoassay solid supports for use in the assay. The
CC solid support is useful for detecting HCV infection in a biological
CC sample. The present sequence is HCV-1 NS3/4a mutant conformational
CC antigen. This sequence is used in the exemplification of the invention
XX
XX

SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
|||||

RESULT 140

AAE21839
ID AAE21839 standard; protein; 686 AA.

AC AAE21839;

DT 16-JUL-2002 (first entry)

DE Hepatitis C virus NS3/4A mutant protein #2.

KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; muten.

OS Hepatitis C virus.

OS Synthetic.

PN WO200214362-A2.

PD 21-FEB-2002.

PF 15-AUG-2001; 2001WO-IB001774.

PR 17-AUG-2000; 2000US-0225767P.

PR 29-AUG-2000; 2000US-0229175P.

PR 03-NOV-2000; 2000US-00705547.

XX (TRIP-) TRIPEP AB.

PA Sallberg M;

PI Sallberg M;

DR WPI; 2002-339446/37.

PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.

PS Claim 19; Page 68-70; 90pp; English.

XX The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
CC cleavage site
XX

SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
|||||

RESULT 141

AAE21845

ID AAE21845 standard; protein; 686 AA.

AC AAE21845;

DT 16-JUL-2002 (first entry)

DE Hepatitis C virus NS3/4A mutant protein #8.

KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; muten.

OS Hepatitis C virus.

OS Synthetic.

PN WO200214362-A2.

PD 21-FEB-2002.

PF 15-AUG-2001; 2001WO-IB001774.

PR 17-AUG-2000; 2000US-0225767P.

PR 29-AUG-2000; 2000US-0229175P.

PR 03-NOV-2000; 2000US-00705547.

XX (TRIP-) TRIPEP AB.

PA Sallberg M;

PI Sallberg M;

DR WPI; 2002-339446/37.

PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.

PS Claim 19; Page 79-81; 90pp; English.

XX The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
CC cleavage site
XX

SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

Db 639 GGVLAALAAAYCLS 651
|||||

RESULT 142

AAE21842

ID AAE21842 standard; protein; 686 AA.

AC AAE21842;

DT 16-JUL-2002 (first entry)

DE Hepatitis C virus NS3/4A mutant protein #5.

KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; muten.

OS Hepatitis C virus.

```

OS Synthetic.
XX WO200214362-A2.
XX PN
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX absence of hepatitis C virus in a subject and for preparing a medicament
XX PD 21-FEB-2002.
XX PF
XX PS 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI
XX PI Sallberg M;
XX WI WPI; 2002-339446/37.
XX DR
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Claim 19; Page 74-75; 90pp; English.
XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX proteins and their corresponding polynucleotides. NS3/4A sequences are
XX useful for identifying the presence or absence HCV in a subject. They are
XX useful for preparing a medicament used for treating or preventing HCV
XX infection. Sequences of the invention are also used as vaccines. The
XX present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
XX cleavage site
XX SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 639 GGVLAALAAAYCLS 651

RESULT 143
AAE21837
ID AAE21837 standard; protein; 686 AA.
AC AAE21837;
XX 16-JUL-2002 (first entry)
XX DE Hepatitis C virus NS3/4A protein.
XX KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX virucide.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN WO200214362-A2.
XX XX
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI
XX PI Sallberg M;
XX WI WPI; 2002-339446/37.
XX DR
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Claim 19; Page 67-68; 90pp; English.
XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX proteins and their corresponding polynucleotides. NS3/4A sequences are
XX useful for identifying the presence or absence HCV in a subject. They are
XX useful for preparing a medicament used for treating or preventing HCV
XX infection. Sequences of the invention are also used as vaccines. The
XX present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
XX cleavage site
XX SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 639 GGVLAALAAAYCLS 651

RESULT 143
AAE21837
ID AAE21837 standard; protein; 686 AA.
AC AAE21837;
XX 16-JUL-2002 (first entry)
XX DE Hepatitis C virus NS3/4A mutant protein #1.
XX KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX virucide; mutant; mutain.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN WO200214362-A2.
XX XX
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI
XX PI Sallberg M;
XX WI WPI; 2002-339446/37.
XX DR
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Claim 19; Page 67-68; 90pp; English.
XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX proteins and their corresponding polynucleotides. NS3/4A sequences are
XX useful for identifying the presence or absence HCV in a subject. They are
XX useful for preparing a medicament used for treating or preventing HCV
XX infection. Sequences of the invention are also used as vaccines. The
XX present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
XX cleavage site
XX SQ Sequence 686 AA;

```

```
Query Match      11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAAAYCLS 30
      |||||
Db      639 GGVLAALAAAYCLS 651

RESULT 145
AAE21840
ID AAE21840 standard; protein; 686 AA.
XX
AC AAE21840;
XX
DT 16-JUL-2002 (first entry)
DE Hepatitis C virus NS3/4A mutant protein #3.
XX
DE Hepatitis C virus NS3/4A mutant protein #3.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX
PI Sallberg M;
XX
DR WPI; 2002-339446/37.
XX
PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
PS Claim 19; Page 70-72; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
CC cleavage site
XX
SQ Sequence 686 AA;

Query Match      11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAAAYCLS 30
      |||||
Db      639 GGVLAALAAAYCLS 651

RESULT 147
AAE21841
ID AAE21841 standard; protein; 686 AA.
XX
AC AAE21841;
XX
DT 16-JUL-2002 (first entry)
DE Hepatitis C virus NS3/4A mutant protein #4.
XX
DE Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
```

```
Query Match      11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAAAYCLS 30
      |||||
Db      639 GGVLAALAAAYCLS 651

RESULT 145
AAE21840
ID AAE21840 standard; protein; 686 AA.
XX
AC AAE21840;
XX
DT 16-JUL-2002 (first entry)
DE Hepatitis C virus NS3/4A mutant protein #3.
XX
DE Hepatitis C virus NS3/4A mutant protein #3.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX
PI Sallberg M;
XX
DR WPI; 2002-339446/37.
XX
PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
PS Claim 19; Page 70-72; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant protein which lacks a proteolytic
CC cleavage site
XX
SQ Sequence 686 AA;

Query Match      11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAAAYCLS 30
      |||||
Db      639 GGVLAALAAAYCLS 651

RESULT 146
AAE21846
ID AAE21846 standard; protein; 686 AA.
XX
AC AAE21846;
XX
```


Db 639 GGVLAALAAAYCLS 651

RESULT 150
AAE19908
ID AAE19908 standard; protein; 686 AA.
XX
AC AAE19908;
XX
XX 18-JUN-2002 (first entry)
XX
DE Hepatitis C virus (HCV) NS3/4A mutant protein #2.
XX
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytotatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
mutant; mutein.
XX
XX Hepatitis C virus.
OS Synthetic.
XX WO200213855-A2.
FN
PD 21-FEB-2002.
XX
PP 15-AUG-2001; 2001WO-IB001808.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
XX (TRIP-) TRIPEP AB.
PA
XX Hepatitis C virus.
OS Synthetic.
XX WO200213855-A2.
FN
PD 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-IB001808.
PF
XX Sallberg M, Hultgren C;
PI WPI; 2002-241837/29.
XX
PT Vaccine compositions for treating and preventing disease, preferably
hepatitis C virus infection, comprises ribavirin and antigen that has
epitope present in hepatitis C virus.
PT Example 6; Page 107-109; 120pp; English.
XX
PS The invention relates to a composition comprising ribavirin and an
antigen preferably non structural 3 protein (NS3)/4A fragment of
hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
sequence. The composition is useful for enhancing an immune response to a
hepatitis C antigen in humans, domestic, sport or pet species and as
vaccines for treating and preventing HCV infections. The composition is
also useful for treating viral, bacterial, fungal diseases and cancer.
The present sequence is HCV NS3/4A mutant protein
Sequence 686 AA;
XX Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX
Oy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
|||||
|||
RESULT 152
AAE19922
ID AAE19922 standard; protein; 686 AA.
XX
AC AAE19922;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis C virus (HCV) NS3/4A mutant protein #6.
XX
KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
cytotatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX WO200213855-A2.
FN
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001808.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX

Db 639 GGVLAALAAAYCLS 651

Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Oy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
|||||
|||
RESULT 151
AAE19919
ID AAE19919 standard; protein; 686 AA.
XX
AC AAE19919;
XX
XX 18-JUN-2002 (first entry)
XX
DE Hepatitis C virus (HCV) NS3/4A mutant protein #3.
XX
KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;

```
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2002-241837/29.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Example 6; Page 113-114; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to a
XX hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A mutant protein
XX
XX Sequence 686 AA;
XX
XX Query Match 11.0%; Score 13; DB 5; Length 686;
XX Best Local Similarity 100.0%; Pred. No. 0.0018;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX Db 639 GGVLAALAAAYCLS 651
XX
XX RESULT 154
XX AAEE19924
XX ID AAEE19924 standard; protein; 686 AA.
XX
XX AC AAEE19924;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Hepatitis C virus (HCV) NS3/4A mutant protein #8.
XX
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX mutant; mutein.
XX
XX OS Hepatitis C virus.
XX
XX OS Synthetic.
XX
XX PN WO200213855-A2.
XX
XX PD 21-FEB-2002.
XX
XX PF 15-AUG-2001; 2001WO-IB001808.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX
XX PR 29-AUG-2000; 2000US-0229175P.
XX
XX PR 03-NOV-2000; 2000US-00705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2002-241837/29.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Example 6; Page 116-118; 120pp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to a
XX hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A mutant protein
XX
XX Sequence 686 AA;
XX
XX Query Match 11.0%; Score 13; DB 5; Length 686;
XX Best Local Similarity 100.0%; Pred. No. 0.0018;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX Db 639 GGVLAALAAAYCLS 651
XX
XX RESULT 153
XX AAEE19907
XX ID AAEE19907 standard; protein; 686 AA.
XX
XX AC AAEE19907;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Hepatitis C virus (HCV) NS3/4A mutant protein #1.
XX
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
XX mutant; mutein.
XX
XX OS Hepatitis C virus.
XX
XX OS Synthetic.
XX
XX PN WO200213855-A2.
XX
XX PD 21-FEB-2002.
XX
XX PF 15-AUG-2001; 2001WO-IB001808.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX
XX PR 29-AUG-2000; 2000US-0229175P.
XX
XX PR 03-NOV-2000; 2000US-00705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2002-241837/29.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Example 6; Page 101-103; 120pp; English.
XX
```


Qy	18	GGVLAALAAAYCLS 30	
Db	639	GGVLAALAAAYCLS 651	
RESULT 155			
AAE19920			
ID	AAE19920	standard; protein; 686 AA.	
XX	AC		
XX	AAE19920;		
XX			
DT	18-JUN-2002	(first entry)	
XX			
DE	Hepatitis C virus (HCV) NS3/4A mutant protein #4.		
XX			
KW	Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;		
KW	cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;		
KW	mutant; muten.		
XX			
OS	Hepatitis C virus.		
OS	Synthetic.		
XX			
FN	WO200213855-A2.		
XX			
PD	21-FEB-2002.		
XX			
PF	15-AUG-2001; 2001WO-IB001808.		
XX			
PR	17-AUG-2000; 2000US-0225767P.		
PR	29-AUG-2000; 2000US-0229175P.		
PR	03-NOV-2000; 2000US-00705547.		
XX			
PA	(TRIP-) TRIPEP AB.		
XX			
PI	Sallberg M, Hultgren C;		
XX			
DR	WPI; 2002-241837/29.		
XX			
PT	Vaccine compositions for treating and preventing disease, preferably		
PT	hepatitis C virus infection, comprises ribavirin and antigen that has		
PT	epitope present in hepatitis C virus.		
XX			
PS	Example 6; Page 109-111; 120pp; English.		
XX			
CC	The invention relates to a composition comprising ribavirin and an		
CC	antigen preferably non structural 3 protein (NS3)/4A fragment of		
CC	hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV		
CC	sequence. The composition is useful for enhancing an immune response to a		
CC	hepatitis C antigen in humans, domestic, sport or pet species and as		
CC	vaccines for treating and preventing HCV infections. The composition is		
CC	also useful for treating viral, bacterial, fungal diseases and cancer.		
CC	The present sequence is HCV NS3/4A mutant protein		
XX			
SQ	Sequence 686 AA;		
Query Match			
Best Local Similarity 11.0%; Score 13; DB 5; Length 686;			
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0			
Qy	18	GGVLAALAAAYCLS 30	
Db	639	GGVLAALAAAYCLS 651	
RESULT 156			
AAE19921			
ID	AAE19921	standard; protein; 686 AA.	
XX			
AC	AAE19921;		
XX			
DT	18-JUN-2002	(first entry)	
XX			

```
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
DR N-PSDB; AAD31767.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Claim 2; Page 95-97; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A protein
XX Sequence 686 AA;
SQ Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLIS 30
Db 639 GGVLAALAAAYCLIS 651
RESULT 159
AAE19925
ID AAE19925 standard; protein; 686 AA.
XX AC AAE19925;
XX DT 18-JUN-2002 (first entry)
XX DE Hepatitis C virus (HCV) NS3/4A mutant protein #9.
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW mutant; muteln.
XX OS Hepatitis C virus.
OS Synthetic.
XX PN WO200213855-A2.
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001808.
XX PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 118-120; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A mutant protein
XX Sequence 686 AA;
SQ Query Match 11.0%; Score 13; DB 5; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLIS 30
Db 639 GGVLAALAAAYCLIS 651
RESULT 158
AAE19923
ID AAE19923 standard; protein; 686 AA.
XX AC AAE19923;
XX DT 18-JUN-2002 (first entry)
XX DE Hepatitis C virus (HCV) NS3/4A mutant protein #7.
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW mutant; muteln.
XX OS Hepatitis C virus.
OS Synthetic.
XX PN WO200213855-A2.
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001808.
XX PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
```



```
CC a biological sample and for treating or detecting non-A, non-B hepatitis
CC (NANB hepatitis). The current sequence is that of the HCV mutant
CC conformational NS3/4a epitope protein of the invention which contains
CC T403P/S404I mutations.
XX
XX
SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 162
ABW00374
ID ABW00374 standard; protein; 686 AA.
XX
XX AC ABW00374;
XX
XX 15-JAN-2004 (first entry)
XX
XX Hepatitis C virus NS3/4A mutant protein #7.
XX
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide; mutant; mutein.
XX
XX Hepatitis C virus.
XX OS Synthetic.
XX
XX US2002136740-A1.
XX
XX 26-SEP-2002.
XX
XX 15-AUG-2001; 2001US-00929955.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALL/) SALLBERG M.
XX PA (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Example 6; Page 82-83; Opp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen, where the antigen is derived from a hepatitis virus. The vaccine
XX is useful in enhancing the immune response to a hepatitis C antigen where
XX the composition is delivered to an animal identified as requiring an
XX enhanced immune response. The vaccine is useful in the treatment and
XX prevention of hepatitis C infection. The present sequence is Hepatitis C
XX virus NS3/4A mutant protein
XX
XX Sequence 686 AA;

Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 164
ABW00370
ID ABW00370 standard; protein; 686 AA.
XX
XX AC ABW00370;
XX
XX 15-JAN-2004 (first entry)
XX
XX Hepatitis C virus NS3/4A mutant protein #3.
XX
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide; mutant; mutein.
XX
XX Hepatitis C virus.
XX OS Synthetic.
XX
```


Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 167
ABW00351
ID ABW00351 standard; protein; 686 AA.
XX AC
XX ABW00351;
XX AC
DT 15-JAN-2004 (first entry)
DE Hepatitis C virus NS3/4A protein.
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
KW virucide.
XX Hepatitis C virus.
OS
XX
PN US2002136740-A1.
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX Sallberg M, Hultgren C;
PI
XX WPI; 2003-764978/72.
DR
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Claim 2; Page 61-63; Opp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A protein
XX
SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 168
ABW00373
ID ABW00373 standard; protein; 686 AA.
XX AC
XX ABW00373;
XX
DT 15-JAN-2004 (first entry)

XX Hepatitis C virus NS3/4A mutant protein #6.
DE Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide; mutant; mutein.
XX Hepatitis C virus.
OS Synthetic.
XX
PN US2002136740-A1.
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX Sallberg M, Hultgren C;
PI
XX WPI; 2003-764978/72.
DR
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Example 6; Page 78-80; Opp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A mutant protein
XX
SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 169
ABW00358
ID ABW00358 standard; protein; 686 AA.
XX AC
XX ABW00358;
XX
DT 15-JAN-2004 (first entry)
DE Hepatitis C virus NS3/4A mutant protein #1.
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
KW virucide; mutant; mutein.
XX Hepatitis C virus.
OS Synthetic.
XX
PN US2002136740-A1.
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.

PR 29-AUG-2000; 2000US-0229175P.
XX (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX Sallberg M, Hultgren C;
PI WPI; 2003-764978/72.
DR Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 67-69; Opp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A mutant protein
XX
SQ Sequence 686 AA;
Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAALAAAYCCLS 30
Db 639 GGVLAAALAAAYCCLS 651
RESULT 170
ABW00372
ID ABW00372 standard; protein; 686 AA.
XX
AC ABW00376;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A mutant protein #9.
XX
KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN US2002136740-A1.
XX
XX 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX
PA (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX
PI Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Example 6; Page 84-85; Opp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A mutant protein
XX
SQ Sequence 686 AA;
Query Match 11.0%; Score 13; DB 7; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAALAAAYCCLS 30
Db 639 GGVLAAALAAAYCCLS 651
RESULT 170
ABW00372
ID ABW00372 standard; protein; 686 AA.
XX
AC ABW00372;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A mutant protein #5.
XX
KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN US2002136740-A1.
XX
XX 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX
PA (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX
PI Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Example 6; Page 76-78; Opp; English.
XX The invention relates to a composition comprising ribavirin and an

```
RESULT 172
ADG47659
ID ADG47659 standard; protein; 686 AA.
XX
AC ADG47659;
XX
DT 11-MAR-2004 (first entry)
XX
DE HCV NS3/4A domain.
XX
KW immunogen; hepatitis C virus; HCV infection; vaccine.
XX
OS Hepatitis C virus.
XX
PN US2003206919-A1.
XX
PD 06-NOV-2003.
XX
PF 26-NOV-2002; 2002US-00307047.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 15-AUG-2001; 2001US-00929955.
XX
PR 15-AUG-2001; 2001US-00930591.
XX
PA (SALL/) SALLBERG M.
XX
PI Sallberg M;
XX
DR WPI; 2004-051480/05.
XX
PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX
PS Example 1; SEQ ID NO 2; 83pp; English.
XX
CC The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX
SQ Sequence 686 AA;
XX
DR WPI; 2004-051480/05.
DR N-ESDB; ADG47658.
XX
PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX
PS Example 1; SEQ ID NO 2; 83pp; English.
XX
CC The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX
SQ Sequence 686 AA;
XX
Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 639 GGVLAALAAAYCLS 651
XX
RESULT 173
ADG47665
ID ADG47665 standard; protein; 686 AA.
XX
AC ADG47665;
XX
DT 11-MAR-2004 (first entry)
XX
DE HCV NS3/4A domain mutant #6.
XX
KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
```

```
PN US2003206919-A1.
XX
PD 06-NOV-2003.
XX
PF 26-NOV-2002; 2002US-00307047.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 15-AUG-2001; 2001US-00929955.
XX
PR 15-AUG-2001; 2001US-00930591.
XX
PA (SALL/) SALLBERG M.
XX
PI Sallberg M;
XX
DR WPI; 2004-051480/05.
XX
PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX
PS Example 1; SEQ ID NO 8; 83pp; English.
XX
CC The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX
SQ Sequence 686 AA;
XX
Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 639 GGVLAALAAAYCLS 651
XX
RESULT 174
ADG47666
ID ADG47666 standard; protein; 686 AA.
XX
AC ADG47666;
XX
DT 11-MAR-2004 (first entry)
XX
DE HCV NS3/4A domain mutant #7.
XX
KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
PN US2003206919-A1.
XX
PD 06-NOV-2003.
XX
PF 26-NOV-2002; 2002US-00307047.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 15-AUG-2001; 2001US-00929955.
XX
PR 15-AUG-2001; 2001US-00930591.
XX
PA (SALL/) SALLBERG M.
XX
PI Sallberg M;
XX
DR WPI; 2004-051480/05.
XX
```


PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.

PS Example 1; SEQ ID NO 9; 83pp; English.

XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.

XX Sequence 686 AA;

SQ Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
639 GGVLAALAAAYCLS 651

RESULT 175

ADG47662
ID ADG47662 standard; protein; 686 AA.

XX AC

ADG47662;

XX DT 11-MAR-2004 (first entry)

XX DE HCV NS3/4A domain mutant #3.

XX immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.

XX Synthetic.

XX Hepatitis C virus.

XX US2003206919-A1.

XX 06-NOV-2003.

XX 26-NOV-2002; 2002US-00307047.

XX 17-AUG-2000; 2000US-0225767P.

XX 29-AUG-2000; 2000US-0229175P.

XX 15-AUG-2001; 2001US-00929955.

XX 15-AUG-2001; 2001US-00930591.

XX (SALL/) SALLBERG M.

XX Sallberg M;

XX WPI; 2004-051480/05.

XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.

PS Example 1; SEQ ID NO 5; 83pp; English.

XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.

XX Sequence 686 AA;

SQ Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
639 GGVLAALAAAYCLS 651

RESULT 176

ADG47693
ID ADG47693 standard; protein; 686 AA.

XX AC

ADG47693;

XX DT 11-MAR-2004 (first entry)

XX DE HCV NS3/4A domain mutant #10.

XX immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.

XX Synthetic.

XX Hepatitis C virus.

XX US2003206919-A1.

XX 06-NOV-2003.

XX 26-NOV-2002; 2002US-00307047.

XX 17-AUG-2000; 2000US-0225767P.

XX 29-AUG-2000; 2000US-0229175P.

XX 15-AUG-2001; 2001US-00929955.

XX 15-AUG-2001; 2001US-00930591.

XX (SALL/) SALLBERG M.

XX Sallberg M;

XX WPI; 2004-051480/05.

XX N-PSDB; ADG47692.

XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.

PS Claim 12; SEQ ID NO 36; 83pp; English.

XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.

XX Sequence 686 AA;

SQ Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
639 GGVLAALAAAYCLS 651

RESULT 177

ADG47660
ID ADG47660 standard; protein; 686 AA.

XX AC

ADG47660;

XX DT 11-MAR-2004 (first entry)

XX DE HCV NS3/4A domain mutant #1.

XX immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX Synthetic.
XX Hepatitis C virus.
XX
XX US2003206919-A1.
XX 06-NOV-2003.
XX
XX 26-NOV-2002; 2002US-00307047.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX (SALL/) SALLBERG M.
XX Sallberg M;
XX WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune response to a hepatitis C antigen comprises specific nucleotide sequences and the amino acid sequences.
XX Example 1; SEQ ID NO 3; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The peptides are useful as immunogens for the treatment and prevention of hepatitis C virus (HCV) infection, in vaccine and immunogen compositions. The nucleic acid and the peptide enhance an immune response to a hepatitis C antigen and are potent immunogens. The present sequence is used in the exemplification of the invention.
XX
XX
SQ Sequence 686 AA;
Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
RESULT 178
ADG47661
ID ADG47661 standard; protein; 686 AA.
AC ADG47661;
XX
XX 11-MAR-2004 (first entry)
XX HCV NS3/4A domain mutant #2.
XX immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX Synthetic.
XX Hepatitis C virus.
XX US2003206919-A1.
XX 06-NOV-2003.
XX 26-NOV-2002; 2002US-00307047.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX

PA (SALL/) SALLBERG M.
XX Sallberg M;
XX WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune response to a hepatitis C antigen comprises specific nucleotide sequences and the amino acid sequences.
XX Example 1; SEQ ID NO 4; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The peptides are useful as immunogens for the treatment and prevention of hepatitis C virus (HCV) infection, in vaccine and immunogen compositions. The nucleic acid and the peptide enhance an immune response to a hepatitis C antigen and are potent immunogens. The present sequence is used in the exemplification of the invention.
XX
XX
SQ Sequence 686 AA;
Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
RESULT 179
ADG47667
ID ADG47667 standard; protein; 686 AA.
AC ADG47667;
XX
XX 11-MAR-2004 (first entry)
XX HCV NS3/4A domain mutant #8.
XX immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX Synthetic.
XX Hepatitis C virus.
XX US2003206919-A1.
XX 06-NOV-2003.
XX 26-NOV-2002; 2002US-00307047.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX (SALL/) SALLBERG M.
XX Sallberg M;
XX WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune response to a hepatitis C antigen comprises specific nucleotide sequences and the amino acid sequences.
XX Example 1; SEQ ID NO 10; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The peptides are useful as immunogens for the treatment and prevention of hepatitis C virus (HCV) infection, in vaccine and immunogen compositions. The nucleic acid and the peptide enhance an immune response to a hepatitis C antigen and are potent immunogens. The present sequence is used in the exemplification of the invention.
XX

CC used in the exemplification of the invention.

XX Sequence 686 AA;

Query Match 11.0%; Score 13; DB 8; Length 686;

Best Local Similarity 100.0%; Pred. No. 0.0018; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

639 GGVLAALAAAYCLS 651

RESULT 180

ADG47664

ID ADG47664 standard; protein; 686 AA.

XX

AC ADG47664;

XX

DT 11-MAR-2004 (first entry)

XX

DE HCV NS3/4A domain mutant #5.

XX

KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.

XX

OS Synthetic.

OS

Hepatitis C virus.

XX

PN US2003206919-A1.

XX

PD 06-NOV-2003.

XX

PF 26-NOV-2002; 2002US-00307047.

XX

PR 17-AUG-2000; 2000US-0225767P.

XX

PR 29-AUG-2000; 2000US-0229175P.

XX

PR 15-AUG-2001; 2001US-00929955.

XX

PR 15-AUG-2001; 2001US-00930591.

XX

PA (SALL/) SALLBERG M.

XX

PI Sallberg M;

XX

DR WPI; 2004-051480/05.

XX

PT New purified or isolated nucleic acid useful for enhancing an immune

response to a hepatitis C antigen comprises specific nucleotide sequences

and the amino acid sequences.

Example 1; SEQ ID NO 7; 83pp; English.

PS The invention relates to a purified or isolated nucleic acid. The

peptides are useful as immunogens for the treatment and prevention of

hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.

The nucleic acid and the peptide enhance an immune response to a

hepatitis C antigen and are potent immunogens. The present sequence is

used in the exemplification of the invention.

SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 8; Length 686;

Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

639 GGVLAALAAAYCLS 651

RESULT 181

ADG47663

ID ADG47663 standard; protein; 686 AA.

XX

AC ADG47663;

XX

DT 11-MAR-2004 (first entry)

XX

DE HCV NS3/4A domain mutant #4.

XX

KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.

XX

OS Synthetic.

OS

Hepatitis C virus.

XX

PN US2003206919-A1.

XX

PD 06-NOV-2003.

XX

PF 26-NOV-2002; 2002US-00307047.

XX

PR 17-AUG-2000; 2000US-0225767P.

XX

PR 29-AUG-2000; 2000US-0229175P.

XX

PR 15-AUG-2001; 2001US-00929955.

XX

PR 15-AUG-2001; 2001US-00930591.

XX

PA (SALL/) SALLBERG M.

XX

PI Sallberg M;

XX

DR WPI; 2004-051480/05.

XX

PT New purified or isolated nucleic acid useful for enhancing an immune

response to a hepatitis C antigen comprises specific nucleotide sequences

and the amino acid sequences.

Example 1; SEQ ID NO 6; 83pp; English.

PS The invention relates to a purified or isolated nucleic acid. The

peptides are useful as immunogens for the treatment and prevention of

hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.

The nucleic acid and the peptide enhance an immune response to a

hepatitis C antigen and are potent immunogens. The present sequence is

used in the exemplification of the invention.

SQ Sequence 686 AA;

Query Match 11.0%; Score 13; DB 8; Length 686;

Best Local Similarity 100.0%; Pred. No. 0.0018;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

639 GGVLAALAAAYCLS 651

RESULT 182

ADG47668

ID ADG47668 standard; protein; 686 AA.

XX

AC ADG47668;

XX

DT 11-MAR-2004 (first entry)

XX

DE HCV NS3/4A domain mutant #9.

XX

KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.

XX

OS Synthetic.

OS

Hepatitis C virus.

XX

PN US2003206919-A1.

XX

PD 06-NOV-2003.

XX

PF 26-NOV-2002; 2002US-00307047.

XX

PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 15-AUG-2001; 2001US-00929955.
PR 15-AUG-2001; 2001US-00930591.
XX (SALL/) SALLBERG M.
XX Sallberg M;
XX WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX Example 1; SEQ ID NO 11; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX Sequence 686 AA;
SQ Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
RESULT 183
ADL66805
ID ADL66805 standard; protein; 686 AA.
XX AC ADL66805;
XX 03-JUN-2004 (first entry)
XX HCV NS3/4a conformational epitope.
XX HCV; NS3/4a conformational epitope; HCV antigen; HCV polyprotein;
KW multiple epitope fusion antigen; MEPA; hepatitis C virus infection.
XX Hepatitis C virus.
XX WO2004021871-A2.
XX 18-MAR-2004.
XX 08-SEP-2003; 2003WO-US028071.
XX 09-SEP-2002; 2002US-0409515P.
XX (CHIR) CHIRON CORP.
XX Arcangel P, Chien D;
XX WPI; 2004-248333/23.
XX N-PSDB; ADL66804.
XX Detecting hepatitis C virus (HCV) infection in a biological sample by
PT detecting complexes formed between the HCV antibody and the antigens from
PT the first region of the HCV polyprotein and the multiple epitope fusion
PT antigen (MEPA).
XX Claim 5; SEQ ID NO 2; 93pp; English.
XX The invention relates to a method of detecting hepatitis C virus (HCV)
CC

CC infection in a biological sample. The method comprises providing an
CC immunoassay solid support comprising HCV antigens bound to it, where the
CC HCV antigens comprise one or more isolated antigens form a first region
CC of the HCV polyprotein, combining a biological sample with the solid
CC support under conditions that allow HCV antibodies, when present in the
CC biological sample, to bind to the one or more HCV antigens, adding to the
CC solid support a detectably labelled HCV multiple epitope fusion antigen
CC (MEPA), where the labelled MEPA comprises at least one epitope from the
CC same region of the HCV polyprotein as the one or more isolated antigens,
CC where the MEPA binds to the bound HCV antibody, and detecting complexes
CC formed between the HCV antibody and the one or more antigens from the
CC first region of the HCV polyprotein and the MEPA, if any, as an
CC indication of HCV infection in the biological sample. The method is
CC useful for detecting hepatitis C virus (HCV) infection in a biological
CC sample. This sequence represents the NS3/4a conformational epitope used
CC in the scope of the invention.
XX Sequence 686 AA;
SQ Query Match 11.0%; Score 13; DB 8; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651
RESULT 184
AAE18688
ID AAE18688 standard; protein; 728 AA.
XX AC AAE18688;
XX 17-MAY-2002 (first entry)
XX NS3/4a mutant conformational antigen.
XX Hepatitis C virus; NS3/4a antigen; HCV infection; mutant; mutuin.
XX Unidentified.
XX Key Location/Qualifiers
FT Misc-difference 182 /note= "Wild type Ser is substituted with Ala"
XX WO200196875-A2.
XX 20-DEC-2001.
XX 14-JUN-2001; 2001WO-US019369.
XX 15-JUN-2000; 2000US-0212082P.
XX 02-APR-2001; 2001US-0280811P.
XX 02-APR-2001; 2001US-0280867P.
XX (CHIR) CHIRON CORP.
XX Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
XX Medina-Selby A;
XX WPI; 2002-179522/23.
XX Immunoassay solid support useful for detecting hepatitis C virus
PT infection in a biological sample, comprises at least one of HCV anti-core
PT antibody and HCV NS3/4a epitope, bound to the support.
XX Disclosure; Fig 3; 87pp; English.
XX The present invention relates to hepatitis C virus (HCV) core antigen and
CC NS (nonstructural) 3/4a antibody combination assay that can detect both
CC HCV antigens and antibodies present in a sample using a single solid
CC matrix as well as immunoassay solid supports for use in the assay. The

CC solid support is useful for detecting HCV infection in a biological
CC sample. The present sequence is NS3/4a mutant conformational antigen.
CC This sequence is used in the exemplification of the invention
XX
SQ Sequence 728 AA;

Query Match 11.0%; Score 13; DB 5; Length 728;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 681 GGVLAALAAAYCLS 693

RESULT 185
ADC06766
ID ADC06766 standard; protein; 728 AA.

XX
AC ADC06766;

XX 18-DEC-2003 (first entry)

XX HCV mutant conformational NS3/4a epitope protein S182A.

XX immunoassay solid support; HCV; NS3/4a; non-structural;
KW non-A, non-B hepatitis; NANB; conformational epitope; mutant; mutein.

XX Synthetic.
OS Hepatitis C virus.

XX
FH Key Location/Qualifiers
FT Misc-difference 182

FT /note= "Wild-type Ser replaced by Ala"

XX US2002192639-A1.

XX 19-DEC-2002.

XX 14-JUN-2001; 2001US-00881239.

XX 15-JUN-2000; 2000US-0212082P.

XX 02-APR-2001; 2001US-0280811P.

XX 02-APR-2001; 2001US-0280867P.

XX (CHIE/) CHIEN D Y.

XX (ARCA/) ARCANDEL P.

XX (TAND/) TANDESKE L.

XX (GEOR/) GEORGE-NASCIMENTO C.

XX (COIT/) COIT D.

XX (MEDI/) MEDINA-SELBY A.

XX Chien DY, Arcangel P, Tandeske L, George-Nascimento C, Coit D;
PI Medina-Selby A;

XX WPI; 2003-644609/61.

XX Immunoassay solid support for detecting hepatitis C virus infection in

XX biological samples, comprises a hepatitis C virus anti-core antibody and
XX an isolated hepatitis C virus NS3/4a epitope bound HCV anti-core
XX antibody.
XX Example 1; Fig 3; 40pp; English.
XX The invention relates to a novel immunoassay solid support comprising at
XX least one hepatitis C virus (HCV) anti-core antibody and at least one
XX isolated HCV NS3/4a (non-structural protein 3/4a) epitope bound thereto.
XX The system of the invention may be useful for detecting HCV infection in
XX a biological sample and for treating or detecting non-A, non-B hepatitis
XX (NANB hepatitis). The current sequence is that of the HCV mutant
XX conformational NS3/4a epitope protein of the invention which contains an
XX S182A mutation.

SQ Sequence 728 AA;

Query Match 11.0%; Score 13; DB 7; Length 728;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 681 GGVLAALAAAYCLS 693

RESULT 186

AAP90146

ID AAP90146 standard; protein; 858 AA.

XX
AC AAP90146;

XX 25-MAR-2003 (revised)

DT 13-NOV-1989 (first entry)

XX ORF extending through hepatitis c virus cDNA in clones 40b, 37b, 35, 36,

DE 81, 32, 33b and 25c.

XX Hepatitis C virus; open reading frame; clone 40b; clone 37b; clone 35;

KW clone 36; clone 81; clone 32; clone 33b; clone 25c; vaccine.

XX Pan troglodytes.

XX GB212511-A.

PN 26-JUL-1989.

XX 18-NOV-1988; 88GB-00027024.

XX 18-NOV-1987; 87US-00122714.

PR 30-DEC-1987; 87US-00139886.

PR 26-FEB-1988; 88US-00161072.

PR 26-OCT-1988; 88US-00263584.

XX (CHIR) CHIRON CORP.

XX Houghton M, Choo QL, Kuo G;

XX WPI; 1989-215054/30.

DR N-PSDB; AAN90310.

XX Hepatitis C virus gene - used for prodn. of polynucleotide probes

PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of

PT infection.

XX Disclosure; Fig 14-1, 14-2, 14-3; 30pp; English.

XX The sequence is encoded by the open reading frame which extends through

CC the hepatitis C virus cDNA of AAN90310. These antigens react with

CC antibodies in patients with non-A non-B hepatitis (NANBH). They can be

CC used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay

CC or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct

CC PR field.)

XX Sequence 858 AA;

Query Match 11.0%; Score 13; DB 1; Length 858;

Best Local Similarity 100.0%; Pred. No. 0.0021;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

DB 399 GGVLAALAAAYCLS 411

RESULT 187

AAP92029

ID AAP92029 standard; protein; 859 AA.

```

XX AC AAP92029;
XX AD
XX DT 09-SEP-2004 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 02-MAR-1990 (first entry)
XX DE HCV protein of cDNA clone inserts.
XX DE Hepatitis C virus (HCV); non-A, non-B hepatitis (NANBH).
XX KW Hepatitis C virus.
XX OS Unidentified.
XX OS EP318216-A.
XX PN
XX PD 31-MAY-1989.
XX PF 18-NOV-1988; 88EP-00310922.
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 06-MAY-1988; 88US-00191263.
XX PR 26-OCT-1988; 88US-00263584.
XX PR 14-NOV-1988; 88US-00271450.
XX PA (CHIR ) CHIRON CORP.
XX PA (CHIR ) CHIRON CORP.
XX PI Houghton M, Choo QL, Kuo G;
XX DR WPI; 1989-159274/22.
XX DR N-PSDB; AAN92085.
XX PI Purified hepatitis C virus - and associated nucleic acids and
XX PT polypeptide(s).
XX PS Claim 13; Fig 14-A, 14-B, 14-C; 139pp; English.
XX CC It is the sequence encoded in the open reading frame of hepatitis C virus
XX CC (HCV) cDNA inserts in clones 40b, 37b, 35, 36, 81, 32, 33b and 25c. It is
XX CC an epitope which could be used as immunoassay reagents and vaccines and
XX CC to generate antibodies useful in diagnosis and passive immunotherapy for
XX CC HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct
XX CC PR field.) (Updated on 25-MAR-2003 to correct PI field.)
XX CC Revised record issued on 09-SEP-2004 : Correction to DE line
XX SQ Sequence 859 AA;
    Query Match 11.0%; Score 13; DB 1; Length 859;
    Best Local Similarity 100.0%; Pred. No. 0.0021;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 399 GGVLAALAAAYCLS 411
    RESULT 188
    AAR33569
    ID AAR33569 standard; protein; 971 AA.
    AC AAR33569;
    XX
    DT 25-MAR-2003 (revised)
    DT 01-JUL-1993 (first entry)
    XX
    XX CKS-HCV antigen fusion protein pHCV-72.
    DE
    DE Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
    KW immunodot assay; Non-A, non-B hepatitis.
    KW

```

```

XX OS Hepatitis C virus.
XX PN WO9304087-A1.
XX PD 04-MAR-1993.
XX PF 21-AUG-1992; 92WO-US007187.
XX PR 21-AUG-1991; 91US-00748566.
XX PA (ABBO ) ABBOTT LAB.
XX PI Desai SM, Casey JM, Rupprecht KR, Devare SG;
XX DR WPI; 1993-093940/11.
XX DE Hepatitis C assay using recombinant C-100 region antigens - for detecting
XX PT antibodies and antigen in body fluids from individuals exposed to
XX PT hepatitis C virus.
XX PS Claim 5; Page 62-66; 206pp; English.
XX CC A specific antigenic region of the HCV genome is expressed as a chimeric
XX CC fusion with E.coli CMP-KDO synthetase (CKS) gene. The fusion protein pHCV
XX CC -72 can be used to detect antibodies and antigen in body fluids from
XX CC individuals exposed to HCV e.g. in confirmatory, competition or
XX CC neutralisation assays. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 971 AA;
    Query Match 11.0%; Score 13; DB 2; Length 971;
    Best Local Similarity 100.0%; Pred. No. 0.0023;
    Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 699 GGVLAALAAAYCLS 711
    RESULT 189
    AAB69027
    ID AAB69027 standard; peptide; 971 AA.
    AC AAB69027;
    XX
    DT 17-APR-2001 (first entry)
    XX
    DE HCV recombinant antigen pHCV-72 amino acid sequence SEQ ID NO:52.
    XX
    KW Hepatitis C virus; HCV; antigen; detection; antibody.
    XX
    OS Hepatitis C virus.
    PN US6172189-B1.
    XX
    PD 09-JAN-2001.
    XX
    PF 02-JUN-1997; 97US-00867611.
    XX
    PR 24-AUG-1990; 90US-00572822.
    PR 07-NOV-1990; 90US-00614069.
    PR 21-AUG-1991; 91US-00748561.
    PR 21-AUG-1991; 91US-00748565.
    PR 21-AUG-1991; 91US-00748566.
    PR 19-NOV-1992; 92US-00989843.
    PR 10-JAN-1994; 94US-00179896.
    PR 01-MAY-1996; 96US-00646757.
    XX
    PA (ABBO ) ABBOTT LAB.
    XX
    PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
    PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

```

XX WPI; 2001-122352/13.
 XX
 PT New recombinant antigens representing distinct antigenic regions of
 PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
 PT antigens in body fluids of individuals exposed to HCV.
 XX
 PS Example 17; Col 187-192; 167pp; English.
 XX
 CC The present invention describes recombinant Hepatitis C virus (HCV)
 CC antigens (I). (I) is useful as a reagent for the detection of antibodies
 CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
 CC uses reliable and efficient reagents and methods to accurately detect the
 CC presence of HCV antibodies in samples obtained from individuals suspected
 CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
 CC AAB69001 to AAB69032 represent sequences used in the exemplification of
 CC the present invention
 XX
 SQ Sequence 971 AA;
 Query Match 11.0%; Score 13; DB 4; Length 971;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 DB 699 GGVLAALAAAYCLS 711
 RESULT 190
 ID ABO1892 standard; protein; 971 AA.
 XX
 AC ABO1892;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DE HCV CKS-C200 recombinant antigen, pHCV-72.
 XX
 KW Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
 KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
 KW fusion protein.
 XX
 OS Chimeric - Hepatitis C virus.
 OS Chimeric - Escherichia coli.
 OS Chimeric - Unidentified.
 XX
 PN US6593083-B1.
 XX
 PD 15-JUL-2003.
 XX
 PF 17-OCT-2000; 2000US-00690359.
 XX
 PR 24-AUG-1990; 90US-00572822.
 PR 07-NOV-1990; 90US-00614069.
 PR 21-AUG-1991; 91US-00748561.
 PR 21-AUG-1991; 91US-00748565.
 PR 21-AUG-1991; 91US-00748566.
 PR 19-NOV-1992; 92US-00989843.
 PR 10-JAN-1994; 94US-00179896.
 PR 01-MAY-1996; 96US-00646757.
 PR 02-JUN-1997; 97US-00867611.
 XX
 PA (DEVA/) DEVARE S G.
 PA (DESA/) DESAI S M.
 PA (CASE/) CASEY J M.
 PA (DALL/) DAILEY S H.
 PA (DAWS/) DAWSON G J.
 PA (GUTI/) GUTIERREZ R A.
 PA (LESN/) LESNIEWSKI R R.
 PA (STEM/) STEWART J L.
 PA (RUPP/) RUPPRECHT K R.
 XX

PI Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 XX WPI; 2003-828264/77.
 DR
 XX Identifying the presence of an antibody in a fluid sample, where the
 PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
 PT antigen by contacting the fluid sample, which may contain a HCV antibody
 PT with at a polypeptide.
 XX
 XX Example 17; Col 187-192; 168pp; English.
 PS
 XX The invention relates to a method for identifying the presence of an
 CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
 CC The method involves providing a fluid sample containing at least one HCV
 CC antibody, contacting the fluid sample with at least one polypeptide or
 CC recombinant fusion protein for complexing the antibody with the
 CC polypeptide or recombinant fusion protein to provide an antibody-
 CC polypeptide complex and detecting the complex. The present sequence is
 CC pHCV-72 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
 CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
 CC enzyme, linker and HCV (non- structural region) NS region. This sequence
 CC is used in the invention
 XX
 SQ Sequence 971 AA;
 Query Match 11.0%; Score 13; DB 7; Length 971;
 Best Local Similarity 100.0%; Pred. No. 0.0023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 DB 699 GGVLAALAAAYCLS 711
 RESULT 191
 AAR33570
 ID AAR33570 standard; protein; 973 AA.
 XX
 AC AAR33570;
 XX
 DT 25-MAR-2003 (revised)
 DT 01-JUL-1993 (first entry)
 XX
 DE CKS-HCV antigen fusion protein pHCV-72'.
 XX
 KW Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
 KW immunodot assay; Non-A, non-B hepatitis.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9304087-A1.
 XX
 PD 04-MAR-1993.
 XX
 PF 21-AUG-1992; 92WO-US007187.
 XX
 PR 21-AUG-1991; 91US-00748566.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Desai SM, Casey JM, Rupprecht KR, Devare SG;
 XX WPI; 1993-093940/11.
 DR
 XX Hepatitis C assay using recombinant C-100 region antigens - for detecting
 PT antibodies and antigen in body fluids from individuals exposed to
 PT hepatitis C virus.
 XX
 PS Claim 6; Page 66-69; 206pp; English.
 XX
 CC A specific antigenic region of the HCV genome is expressed as a chimeric
 CC fusion with E.coli CMP-KDO synthetase (CKS) gene. The fusion protein pHCV

CC -72' can be used to detect antibodies and antigen in body fluids from
CC individuals exposed to HCV e.g. in confirmatory, competition or
CC neutralisation assays. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 973 AA;

Query Match 11.0%; Score 13; DB 2; Length 973;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db |||||

701 GGVLAALAAAYCLS 713

RESULT 192

ID AAB69028 standard; peptide; 973 AA.

XX AAB69028;

DT 17-APR-2001 (first entry)

DE HCV recombinant antigen pHCV-73 amino acid sequence SEQ ID NO:53.

XX Hepatitis C virus; HCV; antigen; detection; antibody.

XX Hepatitis C virus.

XX US6172189-B1.

XX 09-JAN-2001.

XX 02-JUN-1997; 97US-00867611.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX 21-AUG-1991; 91US-00748561.

XX 21-AUG-1991; 91US-00748565.

XX 21-AUG-1991; 91US-00748566.

XX 19-NOV-1992; 92US-00989843.

XX 10-JAN-1994; 94US-00179896.

XX 01-MAY-1996; 96US-00646757.

XX (ABBO) ABBOTT LAB.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2001-122352/13.

XX New recombinant antigens representing distinct antigenic regions of
XX Hepatitis C virus (HCV) genome, useful for detection of antibodies and
XX antigens in body fluids of individuals exposed to HCV.
XX Claim 3; Col 191-198; 167pp; English.

XX The present invention describes recombinant Hepatitis C virus (HCV)

XX antigens (I). (I) is useful as a reagent for the detection of antibodies

XX and antigen in body fluids from individuals exposed to HCV. The HCV assay

XX uses reliable and efficient reagents and methods to accurately detect the

XX presence of HCV antibodies in samples obtained from individuals suspected

XX of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and

XX AAB69001 to AAB69032 represent sequences used in the exemplification of

Db

RESULT 193

ABW01893

ID ABW01893 standard; protein; 973 AA.

XX AC ABW01893;

XX 12-FEB-2004 (first entry)

XX HCV CKS-C200 recombinant antigen, pHCV-73.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;

XX fusion protein.

XX Chimeric - Hepatitis C virus.

XX Chimeric - Escherichia coli.

XX Chimeric - Unidentified.

XX US6593083-B1.

XX 15-JUL-2003.

XX 17-OCT-2000; 2000US-00690359.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX 21-AUG-1991; 91US-00748561.

XX 21-AUG-1991; 91US-00748565.

XX 21-AUG-1991; 91US-00748566.

XX 19-NOV-1992; 92US-00989843.

XX 10-JAN-1994; 94US-00179896.

XX 01-MAY-1996; 96US-00646757.

XX 02-JUN-1997; 97US-00867611.

XX (DEVA/) DEVARE S G.

XX (DESA/) DESAI S M.

XX (CASE/) CASEY J M.

XX (DAILEY/) DAILEY S H.

XX (DAMS/) DAMSON G J.

XX (GUTI/) GUTIERREZ R A.

XX (LESN/) LESNIEWSKI R R.

XX (STEW/) STEWART J L.

XX (RUPP/) RUPPRECHT K R.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

XX antibody is immunologically reactive with a Hepatitis C virus (HCV)

XX antigen by contacting the fluid sample, which may contain a HCV antibody

|||||

701 GGVLAALAAAYCLS 713

RESULT 193

ABW01893

ID ABW01893 standard; protein; 973 AA.

XX AC ABW01893;

XX 12-FEB-2004 (first entry)

XX HCV CKS-C200 recombinant antigen, pHCV-73.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;

XX fusion protein.

XX Chimeric - Hepatitis C virus.

XX Chimeric - Escherichia coli.

XX Chimeric - Unidentified.

XX US6593083-B1.

XX 15-JUL-2003.

XX 17-OCT-2000; 2000US-00690359.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX 21-AUG-1991; 91US-00748561.

XX 21-AUG-1991; 91US-00748565.

XX 21-AUG-1991; 91US-00748566.

XX 19-NOV-1992; 92US-00989843.

XX 10-JAN-1994; 94US-00179896.

XX 01-MAY-1996; 96US-00646757.

XX 02-JUN-1997; 97US-00867611.

XX (DEVA/) DEVARE S G.

XX (DESA/) DESAI S M.

XX (CASE/) CASEY J M.

XX (DAILEY/) DAILEY S H.

XX (DAMS/) DAMSON G J.

XX (GUTI/) GUTIERREZ R A.

XX (LESN/) LESNIEWSKI R R.

XX (STEW/) STEWART J L.

XX (RUPP/) RUPPRECHT K R.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

XX antibody is immunologically reactive with a Hepatitis C virus (HCV)

XX antigen by contacting the fluid sample, which may contain a HCV antibody

|||||

701 GGVLAALAAAYCLS 713

RESULT 193

ABW01893

ID ABW01893 standard; protein; 973 AA.

XX AC ABW01893;

XX 12-FEB-2004 (first entry)

XX HCV CKS-C200 recombinant antigen, pHCV-73.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;

XX fusion protein.

XX Chimeric - Hepatitis C virus.

XX Chimeric - Escherichia coli.

XX Chimeric - Unidentified.

XX US6593083-B1.

XX 15-JUL-2003.

XX 17-OCT-2000; 2000US-00690359.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX 21-AUG-1991; 91US-00748561.

XX 21-AUG-1991; 91US-00748565.

XX 21-AUG-1991; 91US-00748566.

XX 19-NOV-1992; 92US-00989843.

XX 10-JAN-1994; 94US-00179896.

XX 01-MAY-1996; 96US-00646757.

XX 02-JUN-1997; 97US-00867611.

XX (DEVA/) DEVARE S G.

XX (DESA/) DESAI S M.

XX (CASE/) CASEY J M.

XX (DAILEY/) DAILEY S H.

XX (DAMS/) DAMSON G J.

XX (GUTI/) GUTIERREZ R A.

XX (LESN/) LESNIEWSKI R R.

XX (STEW/) STEWART J L.

XX (RUPP/) RUPPRECHT K R.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

XX antibody is immunologically reactive with a Hepatitis C virus (HCV)

XX antigen by contacting the fluid sample, which may contain a HCV antibody

|||||

701 GGVLAALAAAYCLS 713

RESULT 193

ABW01893

ID ABW01893 standard; protein; 973 AA.

XX AC ABW01893;

XX 12-FEB-2004 (first entry)

XX HCV CKS-C200 recombinant antigen, pHCV-73.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;

XX fusion protein.

XX Chimeric - Hepatitis C virus.

XX Chimeric - Escherichia coli.

XX Chimeric - Unidentified.

XX US6593083-B1.

XX 15-JUL-2003.

XX 17-OCT-2000; 2000US-00690359.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX 21-AUG-1991; 91US-00748561.

XX 21-AUG-1991; 91US-00748565.

XX 21-AUG-1991; 91US-00748566.

XX 19-NOV-1992; 92US-00989843.

XX 10-JAN-1994; 94US-00179896.

XX 01-MAY-1996; 96US-00646757.

XX 02-JUN-1997; 97US-00867611.

XX (DEVA/) DEVARE S G.

XX (DESA/) DESAI S M.

XX (CASE/) CASEY J M.

XX (DAILEY/) DAILEY S H.

XX (DAMS/) DAMSON G J.

XX (GUTI/) GUTIERREZ R A.

XX (LESN/) LESNIEWSKI R R.

XX (STEW/) STEWART J L.

XX (RUPP/) RUPPRECHT K R.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

XX antibody is immunologically reactive with a Hepatitis C virus (HCV)

XX antigen by contacting the fluid sample, which may contain a HCV antibody

|||||

701 GGVLAALAAAYCLS 713

RESULT 193

ABW01893

ID ABW01893 standard; protein; 973 AA.

XX AC ABW01893;

XX 12-FEB-2004 (first entry)

XX HCV CKS-C200 recombinant antigen, pHCV-73.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;

XX fusion protein.

XX Chimeric - Hepatitis C virus.

XX Chimeric - Escherichia coli.

XX Chimeric - Unidentified.

XX US6593083-B1.

XX 15-JUL-2003.

XX 17-OCT-2000; 2000US-00690359.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX 21-AUG-1991; 91US-00748561.

XX 21-AUG-1991; 91US-00748565.

XX 21-AUG-1991; 91US-00748566.

XX 19-NOV-1992; 92US-00989843.

XX 10-JAN-1994; 94US-00179896.

XX 01-MAY-1996; 96US-00646757.

XX 02-JUN-1997; 97US-00867611.

XX (DEVA/) DEVARE S G.

XX (DESA/) DESAI S M.

XX (CASE/) CASEY J M.

XX (DAILEY/) DAILEY S H.

XX (DAMS/) DAMSON G J.

XX (GUTI/) GUTIERREZ R A.

XX (LESN/) LESNIEWSKI R R.

XX (STEW/) STEWART J L.

XX (RUPP/) RUPPRECHT K R.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

XX Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

XX antibody is immunologically reactive with a Hepatitis C virus (HCV)

XX antigen by contacting the fluid sample, which may contain a HCV antibody

|||||

701 GGVLAALAAAYCLS 713

RESULT 193

ABW01893

ID ABW01893 standard; protein; 973 AA.

XX AC ABW01893;

XX 12-FEB-2004 (first entry)

XX HCV CKS-C200 recombinant antigen, pHCV-73.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;

XX CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;

XX fusion protein.


```
Query Match      11.0%; Score 13; DB 7; Length 973;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 701 GGVLAALAAAYCLS 713

RESULT 194
AAR33571
ID AAR33571 standard; protein; 992 AA.
XX
AC AAR33571;
XX
DT 25-MAR-2003 (revised)
DT 01-JUL-1993 (first entry)
XX
DE CKS-HCV antigen fusion protein pHCV-18.
XX
KW Hepatitis C virus; C100 antigen; CKS fusion protein; CMP-KDO synthetase;
KW immunodot assay; Non-A, non-B hepatitis.
OS Hepatitis C virus.
XX
PN WO9304087-A1.
XX
PD 04-MAR-1993.
XX
PP 21-AUG-1992; 92WO-US007187.
XX
PR 21-AUG-1991; 91US-00748566.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Desai SM, Casey JM, Rupprecht KR, Devare SG;
XX WPI; 1993-093940/11.
DR
PT Hepatitis C assay using recombinant C-100 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
PT hepatitis C virus.
XX
PS Claim 7; Page 79-73; 206pp; English.
XX
CC A specific antigenic region of the HCV genome is expressed as a chimeric
CC fusion with E.coli CMP-KDO synthetase (CKS) gene. The fusion protein pHCV
CC 18 can be used to detect antibodies and antigen in body fluids from
CC individuals exposed to HCV e.g. in confirmatory, competition or
CC neutralisation assays. The same amino acid sequence (i.e. AAR33571) is
CC described in the specification as identifying the fusion protein of pHCV-
CC 18 and the polypeptide pHCV-205'. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 992 AA;

Query Match      11.0%; Score 13; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 720 GGVLAALAAAYCLS 732

RESULT 195
AAB69029
ID AAB69029 standard; peptide; 992 AA.
XX
AC AAB69029;
XX
DT 17-APR-2001 (first entry)
XX

Query Match      11.0%; Score 13; DB 4; Length 992;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 720 GGVLAALAAAYCLS 732

RESULT 196
ABW01894
ID ABW01894 standard; protein; 992 AA.
XX
AC ABW01894;
XX
DT 12-FEB-2004 (first entry)
XX
DE HCV CKS-C200 recombinant antigen, pHCV-205.
XX
KW Hepatitis C virus; HCV; immunological; CMP-KDO synthetase; CKS;
KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
XX fusion protein.
OS Chimeric - Hepatitis C virus.
OS Chimeric - Escherichia coli.
OS Chimeric - Unidentified.
XX
```

HCV recombinant antigen pHCV-205 amino acid sequence SEQ ID NO:54.
Hepatitis C virus; HCV; antigen; detection; antibody.
Hepatitis C virus.
US6172189-B1.
09-JAN-2001.
02-JUN-1997; 97US-00867611.
24-AUG-1990; 90US-00572822.
07-NOV-1990; 90US-00614069.
21-AUG-1991; 91US-00748561.
21-AUG-1991; 91US-00748565.
21-AUG-1991; 91US-00748566.
19-NOV-1992; 92US-00989843.
10-JAN-1994; 94US-00179896.
01-MAY-1996; 96US-00646757.
(ABBO) ABBOTT LAB.
Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
WPI; 2001-122352/13.
New recombinant antigens representing distinct antigenic regions of
Hepatitis C virus (HCV) genome, useful for detection of antibodies and
antigens in body fluids of individuals exposed to HCV.
Claim 3; Col 197-202; 167pp; English.
The present invention describes recombinant Hepatitis C virus (HCV)
antigens (I). (I) is useful as a reagent for the detection of antibodies
and antigen in body fluids from individuals exposed to HCV. The HCV assay
uses reliable and efficient reagents and methods to accurately detect the
presence of HCV infection. Antigenic regions in samples obtained from suspected
of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
AAB69001 to AAB69032 represent sequences used in the exemplification of
the present invention
Sequence 992 AA;
Query Match 11.0%; Score 13; DB 4; Length 992;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 720 GGVLAALAAAYCLS 732
RESULT 196
ABW01894
ID ABW01894 standard; protein; 992 AA.
XX
AC ABW01894;
XX
DT 12-FEB-2004 (first entry)
XX
DE HCV CKS-C200 recombinant antigen, pHCV-205.
XX
KW Hepatitis C virus; HCV; immunological; CMP-KDO synthetase; CKS;
KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
XX fusion protein.
OS Chimeric - Hepatitis C virus.
OS Chimeric - Escherichia coli.
OS Chimeric - Unidentified.
XX

```
PN U86593083-B1.
XX 15-JUL-2003.
XX 17-OCT-2000; 2000US-00690359.
XX 24-AUG-1990; 90US-00572822.
PR 07-NOV-1990; 90US-00614069.
PR 21-AUG-1991; 91US-00748561.
PR 21-AUG-1991; 91US-00748565.
PR 21-AUG-1991; 91US-00748566.
PR 19-NOV-1992; 92US-00989843.
PR 10-JAN-1994; 94US-00179896.
PR 01-MAY-1996; 96US-00646757.
PR 02-JUN-1997; 97US-00867611.
XX (DEVA/) DEVARE S G.
PA (DESA/) DESAI S M.
PA (CASE/) CASEY J M.
PA (DAIL/) DAILEY S H.
PA (DAWS/) DAWSON G J.
PA (GUTI/) GUTIERREZ R A.
PA (LESN/) LESNIEWSKI R R.
PA (STEM/) STEWART J L.
PA (RUPP/) RUPPRECHT K R.
XX
XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
XX WPI; 2003-828264/77.
XX
XX Identifying the presence of an antibody in a fluid sample, where the
XX antibody is immunologically reactive with a Hepatitis C virus (HCV)
XX antigen by contacting the fluid sample, which may contain a HCV antibody
XX with at a polypeptide.
XX
XX Claim 1; Col 197-202; 168pp; English.
XX
XX The invention relates to a method for identifying the presence of an
XX antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
XX The method involves providing a fluid sample containing at least one HCV
XX antibody, contacting the fluid sample with at least one polypeptide or
XX recombinant fusion protein for complexing the antibody with the
XX polypeptide or recombinant fusion protein to provide an antibody-
XX polypeptide complex and detecting the complex. The present sequence is
XX pHCV-205 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
XX deoxy-manno-octulosonate cytidyl) transferase or CMP-KDO synthase)
XX enzyme, linker and HCV (non- structural region) NS region. This sequence
XX is used in the invention
XX
XX Sequence 992 AA;
XX
XX Query Match 11.0%; Score 13; DB 7; Length 992;
XX Best Local Similarity 100.0%; Pred. No. 0.0024;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLIS 30
XX |||||
XX Db 720 GGVLAALAAAYCLIS 732
XX
XX RESULT 197
XX AAW34481
XX ID AAW34481 standard; protein; 1021 AA.
XX
XX AC AAW34481;
XX
XX DT 25-MAR-2003 (revised)
XX DT 16-MAR-1998 (first entry)
XX
XX HCV antigen combination pSOD/c200/core.
XX
XX PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
KW NS4.
XX
XX Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 1..902
XX /note= "linker"
XX Misc-difference 1..154
XX /note= "hSOD fragment"
XX Misc-difference 155..159
XX /note= "linker"
XX Misc-difference 160..899
XX /note= "c200 (amino acids 1192-1931 of HCV polyprotein)"
XX Misc-difference 903..1021
XX /note= "c22 (amino acids 2-120 of HCV polyprotein)"
XX
XX US5683864-A.
XX
XX 04-NOV-1997.
XX
XX 07-JUL-1992; 92US-00910760.
XX
XX 18-NOV-1987; 87US-00122714.
XX 30-DEC-1987; 87US-00139886.
XX 26-FEB-1988; 88US-00161072.
XX 06-MAY-1988; 88US-00191263.
XX 26-OCT-1988; 88US-00263584.
XX 14-NOV-1988; 88US-00271450.
XX 17-MAR-1989; 89US-00325338.
XX 20-APR-1989; 89US-00341334.
XX 21-APR-1989; 89US-00353896.
XX 18-MAY-1989; 89US-00355002.
XX 04-APR-1990; 90US-00504352.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Kuo G, Houghton M, Choo Q;
XX WPI; 1997-548976/50.
XX N-PSDB; AAT99982.
XX
XX Combination of three hepatitis C virus antigens - used for detection of
XX specific antibodies to diagnose infection.
XX
XX Example 6; Col 59-68; 57pp; English.
XX
XX This sequence represents a Hepatitis c virus (HCV) antigen combination of
XX the invention. The HCV antigen combination comprises an antigen (Agi)
XX comprising the C domain (i.e. amino acids (aa) 1-120 of the HCV
XX polyprotein), or its immunologically reactive fragment containing at
XX least 8 aa. It also comprises two additional antigens from two different
XX polypeptide domains, including at least 8 aa from the NS3, NS4, S or NS5
XX domains of the polyprotein, corresponding, respectively, to aa 1050-1640;
XX 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein. Alternatively,
XX Ag1 contains at least 8 aa from the 1-122 or 9-177 aa regions of the HCV
XX polyprotein. These antigen combinations are used diagnostically to detect
XX anti-HCV antibodies, using any standard immunoassay format. These antigen
XX combinations have a broader range of reactivity with antibodies than any
XX antigen individually. (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 1021 AA;
XX
XX Query Match 11.0%; Score 13; DB 2; Length 1021;
XX Best Local Similarity 100.0%; Pred. No. 0.0024;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLIS 30
XX |||||
XX Db 632 GGVLAALAAAYCLIS 644
```

```

RESULT 198
AAW40039
ID AAW40039 standard; protein; 1021 AA.
XX
AC AAW40039;
XX
DT 26-MAY-1998 (first entry)
XX
DE Fusion protein c200/c22.
XX
KW Hepatitis C virus C domain; HCV; immunological activity; c200/c22;
NS3 domain; NS4 domain; S domain; NS5 domain; fusion protein.
XX
OS Synthetic.
OS Hepatitis virus.
XX
PN US5712087-A.
XX
PD 27-JAN-1998.
XX
PF 12-MAY-1995; 95US-00440519.
XX
PR 04-APR-1990; 90US-00504352.
PR 07-JUL-1992; 92US-00910760.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Kuo G, Houghton M, Choo Q;
XX
DR WPI; 1998-119973/11.
DR N-PSDB; AAV09990.
XX
PT Immunoassays for hepatitis C virus antibodies - using combinations of
PT antigenic fragments of HCV polyprotein.
XX
PS Example 6; Fig 4; 59pp; English.
XX
CC This sequence represents a fusion protein constructed from the hepatitis
CC C virus core domain (which is situated at the carboxy terminus of the
CC fusion protein) and a c200 construct (a fusion of the NS3 and NS3
CC domains). This protein used in the construction of novel combinations of
CC HCV antigens that have a broader range of immunological activity than any
CC single HCV antigen. An example of such an antigen given in this
CC specification comprises a first antigen containing at least 8 amino acids
CC of the C domain of the HCV polyprotein and a second antigen comprising at
CC least 8 amino acids of the NS3 domain, the NS4 domain, the S domain or
CC the NS5 domain of the HCV polyprotein in the form of a fusion protein, a
CC physical mixture or bound to a solid matrix
XX
SQ Sequence 1021 AA;
Query Match 11.0%; Score 13; DB 2; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLIS 30
Db 632 GGVLAALAAAYCLIS 644

RESULT 199
AAE22050
ID AAE22050 standard; protein; 1021 AA.
XX
AC AAE22050;
XX
DT 16-JUL-2002 (first entry)
XX
DE PSOD/c200/core expression plasmid protein.
XX
KW Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
NS4 domain; S domain; NS5 domain; PSOD/c200/core plasmid.
XX

```

```

OS Hepatitis C virus.
OS Unidentified.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT Region 1..154
FT Region /note= "hSOD"
FT Region 155..159
FT Region /note= "Linker region"
FT Region 160..899
FT Region /note= "HCV c200"
FT Region 900..902
FT Region /note= "Linker region"
FT Region 903..1021
FT Region /note= "HCV c22"
XX
PN US6312889-B1.
XX
PD 06-NOV-2001.
XX
PF 12-MAY-1995; 95US-00440549.
XX
PR 04-APR-1990; 90US-00504352.
PR 07-JUL-1992; 92US-00910760.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2002-040268/05.
DR N-PSDB; AAD35044.
XX
PT Combination of hepatitis C viral (HCV) antigens, useful in improved
PT immunoassay for detecting HCV antibodies.
XX
PS Example 6; Fig 4; 58pp; English.
XX
CC The invention relates to combination of hepatitis C viral (HCV) antigens
CC that have a broader range of immunological reactivity than any single HCV
CC antigen. The combinations consist of an antigen from the C domain of the
CC HCV polyprotein, and at least one additional HCV antigen from either the
CC NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
CC the form of fusion protein, a simple physical mixture, or the individual
CC antigens commonly bound to a solid matrix. The combinations of antigens
CC provides broad range immunoassays for anti-HCV antibodies. The invention
CC therefore provides a method for detecting antibodies to HCV in a mammal
CC suspected of containing such antibodies. The present sequence is a
CC protein encoded by PSOD/c200/core expression plasmid DNA containing HCV
CC coding sequence
XX
SQ Sequence 1021 AA;
Query Match 11.0%; Score 13; DB 5; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLIS 30
Db 632 GGVLAALAAAYCLIS 644

RESULT 200
ADY84770
ID ADY84770 standard; protein; 1210 AA.
XX
AC ADY84770;
XX
DT 02-JUN-2005 (first entry)
XX
DE HCV protease NS3.4A-secreted alkaline phosphatase fusion protein.
XX
KW Animal disease model; liver disease; hepatotropic; steatosis;
KW gastrointestinal disease; degeneration; Reye's syndrome; antiemetic;

```

neurological disease; protease; alkaline phosphatase.

XX Hepatitis C virus.
OS Chimeric.
OS Unidentified.

XX Key Location/Qualifiers
XX Region 1..631
FT /label= NS3
FT Region 632..680
FT /label= 4A
FT Region 681..692
FT /label= 4A/4B junction
FT Region 693..1210
FT /label= SEAP

XX WO2005025517-A2.

XX 24-MAR-2005.

XX 13-SEP-2004; 2004WO-US029961.

XX 12-SEP-2003; 2003US-0502742P.

XX 27-OCT-2003; 2003US-0514739P.

XX 01-DEC-2003; 2003US-0526410P.

XX 16-JUL-2004; 2004US-0588909P.

XX (VERT-) VERTEX PHARM INC.

XX Kalkeri G, Kwong A;

XX WPI; 2005-242263/25.

XX N-PSDB; ADY84771.

XX New non-human mammal whose liver comprises a gene system comprising a promoter, and a DNA encoding a protease, a reporter, or protein whose expression causes liver damage, useful as a model for protease activity and for liver damage.

XX Disclosure; Fig 4; 95pp; English.

XX The invention relates to a non-transgenic, non-human animal useful as a model for protease activity and for liver damage, including steatosis and related disorders. A claimed non-human mammal has a liver comprising: (A) a promoter; (B) DNA encoding a protease; and (C) DNA encoding a reporter, where A, B and C are operably linked, and where the presence of reporter activity is indicative of protease activity. The reporter is secreted alkaline phosphatase (SEAP), luciferase, chloramphenicol acetyltransferase, luciferase, beta-galactosidase, green fluorescent protein or horseradish peroxidase, and is detectable in the blood, tissue or serum of the mammal. In another embodiment, the liver of the mammal comprises a gene system comprising an operably linked promoter and DNA encoding a protein whose expression causes liver damage, especially a protease, kinase or esterase. The protease is especially a viral protease such as hepatitis C virus (HCV) NS3.4A protease. The mammal is preferably a mouse, rat or chimpanzee. Methods of producing the mammalian model are claimed. Also claimed are: a hepatocyte or hepatocyte cell line comprising the gene system; methods for identifying compounds that modulate steatosis or which can be used to treat non-alcoholic fatty liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), alcohol steatosis or Reye's syndrome; a method for hepatoprotection in a mammal comprising administering a HCV NS3.4A protease inhibitor; and methods of treating steatosis or fatty liver, alcoholic steatosis, NAFLD, NASH or Reye's syndrome in a mammal by administering a HCV NS3.4A protease inhibitor. The present sequence is that of a fusion protein comprising a HCV protease fused to SEAP via the 4AB junction of HCV.

XX Sequence 1210 AA;

Query Match 11.0%; Score 13; DB 9; Length 1210;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAYCLS 30
Db 639 GGVLAAALAYCLS 651

RESULT 201

ADY84768

ID ADY84768 standard; protein; 1210 AA.

XX ADY84768;

XX 02-JUN-2005 (first entry)

XX HCV protease NS3.4A-secreted alkaline phosphatase fusion protein.

XX Animal disease model; liver disease; hepatotropic; steatosis;
XX gastrointestinal disease; degeneration; Reye's syndrome; antiemetic;
XX neurological disease; protease; alkaline phosphatase; muten.

XX Hepatitis C virus.

XX Chimeric.

XX Synthetic.

XX Unidentified.

XX Key Location/Qualifiers

XX Region 1..631

XX /label= NS3

XX Misc-difference 140 /note= "Wild-type Ser substituted by Ala"

XX Region 632..680

XX /label= 4A

XX Region 681..692

XX /label= 4A/4B junction

XX Region 693..1210

XX /label= SEAP

XX WO2005025517-A2.

XX 24-MAR-2005.

XX 13-SEP-2004; 2004WO-US029961.

XX 12-SEP-2003; 2003US-0502742P.

XX 27-OCT-2003; 2003US-0514739P.

XX 01-DEC-2003; 2003US-0526410P.

XX 16-JUL-2004; 2004US-0588909P.

XX (VERT-) VERTEX PHARM INC.

XX Kalkeri G, Kwong A;

XX WPI; 2005-242263/25.

XX N-PSDB; ADY84769.

XX New non-human mammal whose liver comprises a gene system comprising a promoter, and a DNA encoding a protease, a reporter, or protein whose expression causes liver damage, useful as a model for protease activity and for liver damage.

XX Disclosure; Fig 2; 95pp; English.

XX The invention relates to a non-transgenic, non-human animal useful as a model for protease activity and for liver damage, including steatosis and related disorders. A claimed non-human mammal has a liver comprising: (A) a promoter; (B) DNA encoding a protease; and (C) DNA encoding a reporter, where A, B and C are operably linked, and where the presence of reporter activity is indicative of protease activity. The reporter is secreted alkaline phosphatase (SEAP), luciferase, chloramphenicol acetyltransferase, luciferase, beta-galactosidase, green fluorescent protein or horseradish peroxidase, and is detectable in the blood, tissue or serum of the mammal. In another embodiment, the liver of the mammal comprises a gene system comprising an operably linked promoter and DNA encoding a protein whose expression causes liver damage, especially a

CC protease, kinase or esterase. The protease is especially a viral protease
 CC such as hepatitis C virus (HCV) NS3.4A protease. The mammal is preferably
 CC a mouse, rat or chimpanzee. Methods of producing the mammalian model are
 CC claimed. Also claimed are: a hepatocyte or hepatocyte cell line
 CC comprising the gene system; methods for identifying compounds that
 CC modulate steatosis or which can be used to treat non-alcoholic fatty
 CC liver disease (NAFLD), nonalcoholic steatohepatitis (NASH), alcohol
 CC steatosis or Reye's syndrome; a method for hepatoprotection in a mammal
 CC comprising administering a HCV NS3.4A protease inhibitor; and methods of
 CC treating steatosis or fatty liver, alcoholic steatosis, NAFLD, NASH or
 CC Reye's syndrome in a mammal by administering a HCV NS3.4A protease
 CC inhibitor. The present sequence is that of a fusion protein comprising a
 CC mutant HCV protease fused to SEAP via the 4AB junction of HCV. The Ser-
 CC 139 of NS3 protein in the protease active site is mutated to Ala to
 CC render the protease inactive.

XX Sequence 1210 AA;

Query Match 11.0%; Score 13; DB 9; Length 1210;

Best Local Similarity 100.0%; Pred. No. 0.0028; Mismatches 0; Indels 0; Gaps 0;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

DB 639 GGVLAALAAAYCLS 651

RESULT 202

AAP92041

ID AAP92041 standard; protein; 1766 AA.

XX AC AAP92041;

XX DT 09-SEP-2004 (revised)

DT 25-MAR-2003 (revised)

DT 02-MAR-1990 (first entry)

XX DE Hepatitis C virus (HCV) protein of cDNA clone inserts.

XX KW Hepatitis C virus (HCV); non-A, non-B hepatitis (HABNH).

XX OS Hepatitis C virus.

XX OS Unidentified.

XX PN EP318216-A.

XX PD 31-MAY-1989.

XX PF 18-NOV-1988; 88EP-00310922.

XX PR 18-NOV-1987; 87US-00122714.

PR 30-DEC-1987; 87US-00139886.

PR 26-FEB-1988; 88US-00161072.

PR 06-MAY-1988; 88US-00191263.

PR 26-OCT-1988; 88US-00263584.

PR 14-NOV-1988; 88US-00271450.

XX (CHIR) CHIRON CORP.

PA (CHIR) CHIRON CORP.

XX PI Houghton M, Choo QL, Kuo G;

XX WPI; 1989-159274/22.

DR N-PSDB; AAN92097.

XX PT Purified hepatitis C virus - and associated nucleic acids and

PT polypeptide(s).

XX Claim 13; Fig 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.

XX CC It is the sequence encoded in the open reading frame of hepatitis C virus

CC CDNA inserts in clones 14i.m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81,

CC 32, 33b, 25c, 14c, 8f, a33f, 33g and 39c. It is antigenic and could be

CC used in immunoassay reagents and vaccines and to generate antibodies
 CC useful in diagnosis and passive immunotherapy for HCV infection/non-A,
 CC non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated
 CC on 25-MAR-2003 to correct PI field.)

CC Revised record issued on 09-SEP-2004 : Correction to DE line

XX Sequence 1766 AA;

Query Match 11.0%; Score 13; DB 1; Length 1766;

Best Local Similarity 100.0%; Pred. No. 0.0039;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

DB 948 GGVLAALAAAYCLS 960

RESULT 203

AAB62631

ID AAB62631 standard; protein; 1771 AA.

XX AC AAB62631;

XX DT 23-JUL-2001 (first entry)

XX DE HCV NS35 polypeptide.

XX KW HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;

XX KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;

XX KW immunotherapy; NS35.

XX OS Synthetic.

XX OS Hepatitis C virus.

XX PN WO200138360-A2.

XX PD 31-MAY-2001.

XX PF 22-NOV-2000; 2000WO-US032326.

XX PR 24-NOV-1999; 99US-0167502P.

XX PA (CHIR) CHIRON CORP.

PI Coit D, Medina-Selby A, Selby M, Houghton M;

DR WPI; 2001-343948/36.

DR N-PSDB; AAF83666.

XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as

XX a vaccine against HCV, comprises a polypeptide having a mutation that

XX functionally disrupts the catalytic domain of NS3.

XX Disclosure; Fig 3; 340pp; English.

XX The invention relates to an isolated mutant non-structural (NS) Hepatitis

XX C virus (HCV) polypeptide, comprising a polypeptide having a mutation in

XX the catalytic domain of NS3, where the mutation functionally disrupts the

XX catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a

XX and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide

XX and polynucleotide (preferably DNA or a plasmid) compositions can be used

XX in vaccines against HCV and as diagnostics. The antibodies raised against

XX these polypeptides can also be used as diagnostics, or for passive

XX immunotherapy. The antibodies are also useful for isolating and

XX identifying HCV particles. The present sequence represents the amino acid

XX sequence of a NS35 polypeptide from the plasmid pCMV-NS35

XX Sequence 1771 AA;

Query Match 11.0%; Score 13; DB 4; Length 1771;

Best Local Similarity 100.0%; Pred. No. 0.0039;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      18 GGVLAALAAAYCLS 30
Db      424 GGVLAALAAAYCLS 436

RESULT 204
AAB62634
ID      AAB62634 standard; protein; 1771 AA.
AC      AAB62634;
XX
XX
DT      23-JUL-2001 (first entry)
XX
DE      Amino acid sequence of pd.deltans3NS5.
XX
KW      HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW      catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW      immunotherapy.
XX
OS      Synthetic.
OS      Hepatitis C virus.
XX
PN      WO200138360-A2.
XX
PD      31-MAY-2001.
XX
PF      22-NOV-2000; 2000WO-US032326.
XX
PR      24-NOV-1999; 99US-0167502P.
XX
PA      (CHIR ) CHIRON CORP.
XX
PI      Coit D, Medina-Selby A, Selby M, Houghton M;
XX
DR      WPI; 2001-343948/36.
XX
DR      N-FSDB; AAF83670.
XX
XX
PT      Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
PT      a vaccine against HCV, comprises a polypeptide having a mutation that
PT      functionally disrupts the catalytic domain of NS3.
XX
PS      Claim 32; Fig 11; 340pp; English.
XX
CC      The invention relates to an isolated mutant non-structural (NS) Hepatitis
CC      C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
CC      the catalytic domain of NS3, where the mutation functionally disrupts the
CC      catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC      and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC      and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC      in vaccines against HCV and as diagnostics. The antibodies raised against
CC      these polypeptides can also be used as diagnostics, or for passive
CC      immunotherapy. The antibodies are also useful for isolating and
CC      identifying HCV particles. The present sequence represents the amino acid
CC      sequence of the pd.deltans3NS5
XX
SQ      Sequence 1771 AA;

Query Match      11.0%; Score 13; DB 4; Length 1771;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      424 GGVLAALAAAYCLS 436

RESULT 205
AAB62635
ID      AAB62635 standard; protein; 1771 AA.
AC      AAB62635;
XX
XX
DT      23-JUL-2001 (first entry)
XX
DE      HCV delNS35 polypeptide.
XX
KW      HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW      catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW      immunotherapy; NS35.
XX
OS      Synthetic.
OS      Hepatitis C virus.
XX
PN      WO200138360-A2.
XX

```

```

DT      23-JUL-2001 (first entry)
XX
XX      Amino acid sequence of pd.deltans3NS5.pj.
XX
KW      HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW      catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW      immunotherapy.
XX
OS      Synthetic.
OS      Hepatitis C virus.
XX
PN      WO200138360-A2.
XX
PD      31-MAY-2001.
XX
PF      22-NOV-2000; 2000WO-US032326.
XX
PR      24-NOV-1999; 99US-0167502P.
XX
PA      (CHIR ) CHIRON CORP.
XX
PI      Coit D, Medina-Selby A, Selby M, Houghton M;
XX
DR      WPI; 2001-343948/36.
XX
DR      N-FSDB; AAF83671.
XX
XX
PT      Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
PT      a vaccine against HCV, comprises a polypeptide having a mutation that
PT      functionally disrupts the catalytic domain of NS3.
XX
PS      Example 1; Fig 14; 340pp; English.
XX
CC      The invention relates to an isolated mutant non-structural (NS) Hepatitis
CC      C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
CC      the catalytic domain of NS3, where the mutation functionally disrupts the
CC      catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC      and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC      and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC      in vaccines against HCV and as diagnostics. The antibodies raised against
CC      these polypeptides can also be used as diagnostics, or for passive
CC      immunotherapy. The antibodies are also useful for isolating and
CC      identifying HCV particles. The present sequence represents the amino acid
CC      sequence of the pd.deltans3NS5.pj
XX
SQ      Sequence 1771 AA;

Query Match      11.0%; Score 13; DB 4; Length 1771;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      424 GGVLAALAAAYCLS 436

RESULT 206
AAB62632
ID      AAB62632 standard; protein; 1771 AA.
XX
AC      AAB62632;
XX
DT      23-JUL-2001 (first entry)
XX
DE      HCV delNS35 polypeptide.
XX
KW      HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW      catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW      immunotherapy; NS35.
XX
OS      Synthetic.
OS      Hepatitis C virus.
XX
PN      WO200138360-A2.
XX

```

XX 31-MAY-2001.
PD
XX
PF 22-NOV-2000; 2000WO-US032326.
XX
XX 24-NOV-1999; 99US-0167502P.
PR
XX (CHIR) CHIRON CORP.
PA
XX Coit D, Medina-Selby A, Selby M, Houghton M;
PI WPI; 2001-343948/36.
XX
XX N-PSDB; AAF83667.
DR
XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
PT a vaccine against HCV, comprises a polypeptide having a mutation that
PT functionally disrupts the catalytic domain of NS3.
XX
XX Example 1; Fig 5; 340pp; English.
PS
XX The invention relates to an isolated mutant non-structural (NS) Hepatitis
CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
CC the catalytic domain of NS3, where the mutation functionally disrupts the
CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC in vaccines against HCV and as diagnostics. The antibodies raised against
CC these polypeptides can also be used as diagnostics, or for passive
CC immunotherapy. The antibodies are also useful for isolating and
CC identifying HCV particles. The present sequence represents the amino acid
CC sequence of a delNS35 polypeptide from the plasmid pCMV-delNS35
XX
XX Sequence 1771 AA;
Query Match 11.0%; Score 13; DB 4; Length 1771;
Best Local Similarity 100.0%; Pred. No. 0.0039; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCCLS 30
Db 424 GGVLAALAAAYCCLS 436
|||||
RESULT 207
AAB62636
ID AAB62636 standard; protein; 1892 AA.
XX
XX AAB62636;
AC
XX 23-JUL-2001 (first entry)
DT
XX Amino acid sequence of pd.deltans3NS5.pj.core121.
DE
XX HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
KW immunotherapy.
XX
XX Synthetic.
OS
XX Hepatitis C virus.
OS
XX WO200138360-A2.
FN
XX 31-MAY-2001.
PD
XX
XX 22-NOV-2000; 2000WO-US032326.
PF
XX 24-NOV-1999; 99US-0167502P.
PR
XX (CHIR) CHIRON CORP.
PA
XX Coit D, Medina-Selby A, Selby M, Houghton M;
PI WPI; 2001-343948/36.
XX
DR

DR N-PSDB; AAF83672.
XX
XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
PT a vaccine against HCV, comprises a polypeptide having a mutation that
PT functionally disrupts the catalytic domain of NS3.
XX
XX Example 1; Fig 17; 340pp; English.
PS
XX The invention relates to an isolated mutant non-structural (NS) Hepatitis
CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
CC the catalytic domain of NS3, where the mutation functionally disrupts the
CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
CC and polynucleotide (preferably DNA or a plasmid) compositions can be used
CC in vaccines against HCV and as diagnostics. The antibodies raised against
CC these polypeptides can also be used as diagnostics, or for passive
CC immunotherapy. The antibodies are also useful for isolating and
CC identifying HCV particles. The present sequence represents the amino acid
CC sequence of the pd.deltans3NS5.pj.core121
XX
XX Sequence 1892 AA;
Query Match 11.0%; Score 13; DB 4; Length 1892;
Best Local Similarity 100.0%; Pred. No. 0.0041; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCCLS 30
Db 424 GGVLAALAAAYCCLS 436
|||||
RESULT 208
ADI34636
ID ADI34636 standard; protein; 1892 AA.
XX
XX ADI34636;
AC
XX 22-APR-2004 (first entry)
DT
XX HCVmodified fusion protein.
DE
XX NS3; HCV; protease; polypeptide; antiinflammatory; hepatotropic;
KW virucide; gene therapy; vaccine; immunogenic; cellular immune response;
KW fusion protein.
XX
XX Hepatitis C virus.
OS
XX Synthetic.
OS
XX WO2004005473-A2.
FN
XX 15-JAN-2004.
PD
XX 02-JUL-2003; 2003WO-US020996.
PF
XX 02-JUL-2002; 2002US-0393694P.
PR
XX 08-JUL-2002; 2002US-0394510P.
PR
XX (CHIR) CHIRON CORP.
PA
XX Houghton M;
PI
XX WPI; 2004-091351/09.
DR
XX N-PSDB; ADI34635.
DR
XX New hepatitis C virus (HCV) fusion proteins with modified NS3 domains,
PT useful for manufacturing a medicament for stimulating a cellular immune
PT response, particularly for preventing or treating HCV infection, in
PT mammals.
XX
XX Disclosure; SEQ ID NO 6; 86pp; English.
PS
XX The invention relates to an immunogenic fusion protein comprising a
CC modified NS3 polypeptide containing an amino acid substitution to the

CC hepatitis C virus (HCV) NS3 region, so that protease activity is
 CC inhibited, and a polypeptide derived from a region of the HCV polyprotein
 CC other than the NS3 region. The modification comprises a substitution of
 CC an amino acid corresponding to His-1083, Asp-1105 and/or Ser-1165,
 CC numbered relative to the full-length HCV-1 polyprotein. The immunogenic
 CC fusion protein or the modified NS3 polypeptide are useful in
 CC manufacturing a medicament or a composition for stimulating a cellular
 CC immune response in a vertebrate subject (claimed). These may be used in
 CC identifying epitopes of HCV polypeptides associated with a strong
 CC cytotoxic T-lymphocyte (CTL) response, or in preventing or treating HCV
 CC infection in mammals, including humans. The present sequence represents a
 CC representative HCV modified fusion protein, with the NS3 protease domain
 CC deleted from the N-terminus and including the 1-121 of Core on the C-
 CC terminus.
 XX
 SQ Sequence 1892 AA;

Query Match 11.0%; Score 13; DB 8; Length 1892;
 Best Local Similarity 100.0%; Pred. No. 0.0041;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||
 Db 424 GGVLAALAAAYCLS 436

RESULT 209

ADO00774
 ID ADO00774 standard; protein; 1892 AA.

XX ADO00774;

DT 29-JUL-2004 (first entry)

DE HCV NS345Core fusion protein SEQ ID NO:8.

XX fusion protein; hepatitis C virus; HCV; NS3; NS4; core polypeptide;
 KW polypeptide; HCV-1; antiviral; vaccine; T cell activation;
 KW NS345Core fusion protein; NS5.

XX Hepatitis C virus.
 OS Synthetic.

XX WO2004039950-A2.

XX 13-MAY-2004.

PF 24-OCT-2003; 2003WO-US033610.

XX 25-OCT-2002; 2002US-00281341.

XX (CHIR) CHIRON CORP.

XX Houghton M, Coates S, Selby M, Paliard X;

XX WPI; 2004-376177/35.

DR N-PSDB; ADO00773.

XX New fusion protein comprises hepatitis C virus (HCV) polypeptides, which
 PT consist of an NS3, an NS4, an NS4 and a core polypeptide of HCV, useful
 PT for stimulating immune response, e.g. activating T cells of a vertebrate.

PS Claim 6; SEQ ID NO 8; 136pp; English.

XX The present invention describes a fusion protein comprising hepatitis C
 CC virus (HCV) polypeptides, where the HCV polypeptides consist of an NS3,
 CC an NS4, an NS5 and a core polypeptide of HCV, and the core polypeptide
 CC consist of amino acids 1-121 of the HCV polyprotein, numbered relative to
 CC the full-length HCV-1 polyprotein. Also described: (1) an isolated and
 CC purified polynucleotide that encodes the fusion protein; (2) a
 CC composition comprising the fusion protein or the isolated and purified
 CC polynucleotide and a pharmaceutical excipient; and (3) activating T cells
 CC of a vertebrate subject which recognise an epitope of an HCV polypeptide.

CC The fusion protein has antiviral activity, and can be used in vaccines.
 CC The composition is useful for activating T cells of a vertebrate subject
 CC which recognise an epitope of an of the NS3, NS4, NS5 and/or core
 CC polypeptides. It is also useful for the manufacturing of a medicament for
 CC activating T cells. The present sequence represents an HCV NS345Core
 CC fusion protein amino acid sequence, which is used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 1892 AA;

Query Match 11.0%; Score 13; DB 8; Length 1892;
 Best Local Similarity 100.0%; Pred. No. 0.0041;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||
 Db 424 GGVLAALAAAYCLS 436

RESULT 210

AAB62638
 ID AAB62638 standard; protein; 1911 AA.

XX AAB62638;

DT 23-JUL-2001 (first entry)

DE Amino acid sequence of pd.deltans3NS5.pj.core140.

XX HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;
 KW catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;
 KW immunotherapy.

XX Synthetic.

OS Hepatitis C virus.

XX WO200138360-A2.

XX 31-MAY-2001.

XX 22-NOV-2000; 2000WO-US032326.

PR 24-NOV-1999; 99US-0167502P.

XX (CHIR) CHIRON CORP.

XX Coit D, Medina-Selby A, Selby M, Houghton M;

XX WPI; 2001-343948/36.

DR N-PSDB; AAF83674.

XX Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as
 PT a vaccine against HCV, comprises a polypeptide having a mutation that
 PT functionally disrupts the catalytic domain of NS3.

PS Example 1; Fig 21; 340pp; English.

XX The invention relates to an isolated mutant non-structural (NS) Hepatitis
 CC C virus (HCV) polypeptide, comprising a polypeptide having a mutation in
 CC the catalytic domain of NS3, where the mutation functionally disrupts the
 CC catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a
 CC and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide
 CC and polynucleotide (preferably DNA or a plasmid) compositions can be used
 CC in vaccines against HCV and as diagnostics. The antibodies raised against
 CC these polypeptides can also be used as diagnostics, or for passive
 CC immunotherapy. The antibodies are also useful for isolating and
 CC identifying HCV particles. The present sequence represents the amino acid
 CC sequence of the pd.deltans3NS5.pj.core140

XX Sequence 1911 AA;

Query Match 11.0%; Score 13; DB 4; Length 1911;
 Best Local Similarity 100.0%; Pred. No. 0.0041;

Matches	13;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	18	GGVLAALAAYCLS	30						
Db	424	GGVLAALAAYCLS	436						
RESULT 211									
AAB62639		standard; protein; 1921 AA.							
AC	XX								
AC	AAB62639;								
DT	23-JUL-2001	(first entry)							
XX	XX								
DE	XX	Amino acid sequence of pd.deltans3NS5.pj.core150.							
KW	XX	HCV; mutant; non-structural protein; NS; hepatitis C virus; mutation;							
KW	KW	catalytic domain; NS3; NS4; NS5; antiviral; vaccine; immunostimulant;							
KW	XX	immunotherapy.							
OS	XX	Synthetic.							
OS	XX	Hepatitis C virus.							
PN	XX	WO200138360-A2.							
XX	XX	31-MAY-2001.							
XX	XX	22-NOV-2000; 2000WO-US032326.							
PF	XX	24-NOV-1999; 99US-0167502P.							
PR	XX	(CHIR) CHIRON CORP.							
PA	XX	Colt D, Medina-Selby A, Selby M, Houghton M;							
PI	XX	WPI; 2001-343948/36.							
DR	XX	N-FSDB; AAF83675.							
DR	XX	Mutant non-structural (NS) Hepatitis C virus (HCV) polypeptide, useful as							
PT	XX	a vaccine against HCV, comprises a polypeptide having a mutation that							
PT	XX	functionally disrupts the catalytic domain of NS3.							
XX	XX	Example 1; Fig 22; 340pp; English.							
XX	XX	The invention relates to an isolated mutant non-structural (NS) Hepatitis							
CC	XX	C virus (HCV) polypeptide, comprising a polypeptide having a mutation in							
CC	XX	the catalytic domain of NS3, where the mutation functionally disrupts the							
CC	XX	catalytic domain. The NS mutant polypeptides can include NS3, NS4 (NS4a							
CC	XX	and NS4b) NS5 (NS5a and NS5b) or portions thereof. The HCV polypeptide							
CC	XX	and polynucleotide (preferably DNA or a plasmid) compositions can be used							
CC	XX	in vaccines against HCV and as diagnostics. The antibodies raised against							
CC	XX	these polypeptides can also be used as diagnostics, or for passive							
CC	XX	immunotherapy. The antibodies are also useful for isolating and							
CC	XX	identifying HCV particles. The present sequence represents the amino acid							
CC	XX	sequence of the pd.deltans3NS5.pj.core150							
SQ	XX	Sequence 1921 AA;							
Query Match 11.0%; Score 13; DB 4; Length 1921;									
Best Local Similarity 100.0%; Pred. No. 0.0041;									
Matches	13;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	18	GGVLAALAAYCLS	30						
Db	424	GGVLAALAAYCLS	436						
RESULT 212									
AAB62637		standard; protein; 1944 AA.							
ID	XX								
XX	XX								
AC	AAB62637;								


```
XX DE HCV HepC cassette C.
XX DE Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX DE viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia.
XX OS Hepatitis C virus.
XX PN WO200190197-A1.
XX PD 29-NOV-2001.
XX XX
XX PF 25-MAY-2001; 2001WO-AU000622.
XX XX
XX PR 26-MAY-2000; 2000AU-00007761.
XX XX
XX PA (AUSU ) UNIV AUSTRALIAN NAT.
XX PI Thomson SA, Ramshaw IA;
XX XX
XX DR WPI; 2002-147575/19.
XX DR N-PSDB; ABK36640.
XX XX
XX PT New synthetic polypeptides having several different segments of at least
XX PT one parent polypeptide linked together differently compared to the
XX PT linkage in the parent polypeptide, for inducing immune response against a
XX PT pathogen or cancer.
XX PS Example 2; SEQ ID NO 816; 364pp; English.
XX XX
XX CC The invention relates to a new synthetic polypeptide (I) comprising
XX CC several different segments of at least one parent polypeptide linked
XX CC together in a different relationship relative to their linkage in the
XX CC parent polypeptide to impede, abrogate or otherwise alter at least one
XX CC function associated with the parent polypeptide and for inducing an
XX CC immune response against a pathogen or cancer. Also included are a
XX CC synthetic polynucleotide encoding and a computer system for designing the
XX CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
XX CC are referred to as a Savine. The synthetic polypeptide is useful for
XX CC modulating immune responses preferably directed against a pathogen or a
XX CC cancer (e.g., cancers of the lung, breast, ovary, cervix, bone liver,
XX CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
XX CC oesophagus, brain, testicle, uterus), as potentiating agents.
XX CC Compositions comprising the polypeptide may be used in the treatment or
XX CC prophylaxis against viral (such as infections caused by HIV (human
XX CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
XX CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
XX CC a cassette protein consisting of several peptides derived from a parent
XX CC protein. One or more cassettes are used to construct a savine of the
XX CC invention.
XX SQ Sequence 1997 AA;

Query Match 11.0%; Score 13; DB 5; Length 1997;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
| | | | | | | | | |
Db 436 GGVLAALAAAYCLIS 448

RESULT 216
AAU84800
ID AAU84800 standard; protein; 2011 AA.
XX
AC AAU84800;
```

```
XX DT 08-MAY-2002 (first entry)
XX XX
XX DE HCV HepC cassette A.
XX XX
XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX KW viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia.
XX OS Hepatitis C virus.
XX PN WO200190197-A1.
XX PD 29-NOV-2001.
XX XX
XX PF 25-MAY-2001; 2001WO-AU000622.
XX XX
XX PR 26-MAY-2000; 2000AU-00007761.
XX XX
XX PA (AUSU ) UNIV AUSTRALIAN NAT.
XX PI Thomson SA, Ramshaw IA;
XX XX
XX DR WPI; 2002-147575/19.
XX DR N-PSDB; ABK36638.
XX XX
XX PT New synthetic polypeptides having several different segments of at least
XX PT one parent polypeptide linked together differently compared to the
XX PT linkage in the parent polypeptide, for inducing immune response against a
XX PT pathogen or cancer.
XX PS Example 2; Fig 26; 364pp; English.
XX XX
XX CC The invention relates to a new synthetic polypeptide (I) comprising
XX CC several different segments of at least one parent polypeptide linked
XX CC together in a different relationship relative to their linkage in the
XX CC parent polypeptide to impede, abrogate or otherwise alter at least one
XX CC function associated with the parent polypeptide and for inducing an
XX CC immune response against a pathogen or cancer. Also included are a
XX CC synthetic polynucleotide encoding and a computer system for designing the
XX CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
XX CC are referred to as a Savine. The synthetic polypeptide is useful for
XX CC modulating immune responses preferably directed against a pathogen or a
XX CC cancer (e.g., cancers of the lung, breast, ovary, cervix, colon, head
XX CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
XX CC oesophagus, brain, testicle, uterus), as potentiating agents.
XX CC Compositions comprising the polypeptide may be used in the treatment or
XX CC prophylaxis against viral (such as infections caused by HIV (human
XX CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
XX CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
XX CC a cassette protein consisting of several peptides derived from a parent
XX CC protein. One or more cassettes are used to construct a savine of the
XX CC invention.
XX SQ Sequence 2011 AA;

Query Match 11.0%; Score 13; DB 5; Length 2011;
Best Local Similarity 100.0%; Pred. No. 0.0043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
| | | | | | | | | |
Db 1592 GGVLAALAAAYCLIS 1604

RESULT 217
AAO26783
ID AAO26783 standard; protein; 2202 AA.
```

```
XX AAO26783;
AC
XX 03-APR-2003 (first entry)
DT
XX
XX Protein derived from Hepatitis C virus DNA, SEQ ID NO 1.
DE
XX Virucide; hepatotropic; antiinflammatory; prenylation; hepatitis C virus;
KW HCV; NS5A protein; anti-HCV agent; HCV infection.
XX
XX Hepatitis C virus.
OS
XX
FH Key Location/Qualifiers
FT Misc-difference 359 /label= His
FT /note= "Residue encoded by CCG"
FT Misc-difference 360 /label= Ala
FT /note= "Residue encoded by CTG"
FT Misc-difference 361 /label= Val
FT /note= "Residue encoded by TTG"
FT Misc-difference 362 /label= Gly
FT /note= "Residue encoded by TGC"
FT Misc-difference 363 /label= Leu
FT /note= "Residue encoded by CCC"
FT Misc-difference 364 /label= Phe
FT /note= "Residue encoded by GCG"
FT Misc-difference 365 /label= Arg
FT /note= "Residue encoded by GGA"
FT Misc-difference 366 /label= Ala
FT /note= "Residue encoded by CAC"
FT Misc-difference 369 /label= Cys
FT /note= "Residue encoded by GGC"
FT Misc-difference 370 /label= Thr
FT /note= "Residue encoded by CTA"
FT Misc-difference 371 /label= Arg
FT /note= "Residue encoded by TTC"
FT Misc-difference 372 /label= Gly
FT /note= "Residue encoded by AGG"
FT Misc-difference 373 /label= Pro
FT /note= "Residue encoded by GCC"
FT Misc-difference 374 /label= Leu
FT /note= "Residue encoded by GCG"
FT Misc-difference 375 /label= Leu
FT /note= "Residue encoded by GTG"
FT Misc-difference 377 /label= Pro
FT /note= "Residue encoded by ACC"
FT Misc-difference 378 /label= Ala
FT /note= "Residue encoded by CGT"
XX
PN FR2824072-A1.
XX
XX 31-OCT-2002.
XX
XX 27-APR-2001; 2001PR-00005732.
XX
XX 27-APR-2001; 2001PR-00005732.
XX
```

```
PA (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Wychowski C, Duverlie G, Dubuisson J, Pillez A;
PI
XX WPI; 2003-142456/14.
XX
XX N-PSDB; AAL54423.
DR
XX
XX Using cells that can prenylate proteins for replication and production of
PT hepatitis C virus (HCV), useful for screening compounds for anti-HCV
PT activity.
PT
XX
XX Disclosure; Fig 1; 85pp; French.
PS
XX
XX The invention relates to cells which are able to cause prenylation of
CC proteins encoded by the genome of hepatitis C virus (HCV), such as NS5A
CC protein, for replication, and optionally production, of HCV or its
CC derived viable mutants, in an appropriate culture medium. The cells are
CC used to produce HCV particles, and to screen for anti-HCV agents. This
CC inhibitors of prenylation are useful for treating HCV infection. This
CC sequence represents an HCV protein relating to the invention
CC
XX Sequence 2202 AA;
SQ
XX
XX Query Match 11.0%; Score 13; DB 6; Length 2202;
XX Best Local Similarity 100.0%; Pred. No. 0.0046;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLS 30
XX |||||
XX DB 855 GGVLAALAAAYCLS 867
XX
XX RESULT 218
XX AAP90164
XX ID AAP90164 standard; protein; 2261 AA.
XX
XX AC AAP90164;
XX
XX DT 25-MAR-2003 (revised)
XX DT 01-NOV-1989 (first entry)
XX
XX DE Peptide encoded by composite hepatitis C virus cDNA.
XX
XX KW Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.
XX
XX OS Pan troglodytes.
XX
XX PN GB2212511-A.
XX
XX PD 26-JUL-1989.
XX
XX PF 18-NOV-1988; 88GB-00027024.
XX
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 26-OCT-1988; 88US-00263584.
XX
XX (CHIR ) CHIRON CORP.
PA
XX
XX Houghton M, Choo QL, Kuo G;
PI
XX WPI; 1989-215054/30.
XX
XX N-PSDB; AAN90331.
DR
XX
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
PT
XX
XX Disclosure; Fig 32; 30pp; English.
PS
XX
XX The sequence is the peptide encoded by the composite hepatitis C virus
CC (HCV) cDNA of AAN90331. The polypeptides are used to diagnose HCV-induced
CC
```

CC NANBH, to raise antibodies for immunoassay or treatment, or to produce
CC vaccines. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 2261 AA;

Query Match 11.0%; Score 13; DB 1; Length 2261;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 1039 GGVLAALAAAYCLS 1051

RESULT 219
AAP92047
ID AAP92047 standard; protein; 2301 AA.
XX
AC AAP92047;
XX
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE HCV protein of cDNA inserts in clones 12f through 15e.
XX
KW Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH).
XX
OS Hepatitis C virus.
OS Unidentified.
XX
PN EP318216-A.
XX
PD 31-MAY-1989.
XX
PF 18-NOV-1988; 88EP-00310922.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
PA (CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-159274/22.
DR N-PSDB; AAN92103.
XX
PT Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
PS Claim 13; Fig 32-1-32-7; 139pp; English.
XX
CC It is the sequence encoded in the open reading frame of hepatitis C virus
CC (HCV) cDNA inserts in clones 12f through 15e. It is antigenic and could
CC be used in immunoassay reagents and vaccines and to generate antibodies
CC useful in diagnosis and passive immunotherapy for HCV infection/non-A,
CC non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated
CC on 25-MAR-2003 to correct PI field.)
CC
CC Revised record issued on 09-SEP-2004 : Correction to DB line
XX
SQ Sequence 2301 AA;

Query Match 11.0%; Score 13; DB 1; Length 2301;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

Db 1039 GGVLAALAAAYCLS 1051

RESULT 220
AAR25135
ID AAR25135 standard; protein; 2435 AA.
XX
AC AAR25135;
XX
DT 23-DEC-1992 (first entry)
XX
DE HCV polypeptide 1.
XX
KW Hepatitis C virus; blood transfusion.
XX
OS Synthetic.
XX
PN JP04159298-A.
XX
PD 02-JUN-1992.
XX
PF 19-OCT-1990; 90JP-00282431.
XX
PR 19-OCT-1990; 90JP-00282431.
XX
PA (OLYU) OLYMPUS OPTICAL CO LTD.
XX
DR WPI; 1992-231947/28.
XX
PT New peptides acting as antigenic analogues of human hepatitis C virus -
PT useful for detecting HCV antibody positive patients.
XX
PS Claim 1; Page 1; 14pp; Japanese.
XX
CC The sequences given in AAR25135-36 are peptides from the hepatitis C
CC virus (HCV) which are recognised by the peptide sequences given in
CC AAR25130-24. These antigenic peptides can be used on their own or as a
CC mixture two different peptides. Using these peptides, HCV antibody
CC positive patients can be detected and hepatitis caused by blood
CC transfusion can be prevented
XX
SQ Sequence 2435 AA;

Query Match 11.0%; Score 13; DB 2; Length 2435;
Best Local Similarity 100.0%; Pred. No. 0.0051;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 1214 GGVLAALAAAYCLS 1226

RESULT 221
AAP92050
ID AAP92050 standard; protein; 2436 AA.
XX
AC AAP92050;
XX
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE HCV protein of the cDNA inserts in clones K9-1 through 15e.
XX
KW Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH).
XX
OS Hepatitis C virus.
OS Unidentified.
XX
PN EP318216-A.
XX
PD 31-MAY-1989.

```
XX PF 18-NOV-1988; 88EP-00310922.
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 06-MAY-1988; 88US-00191263.
XX PR 26-OCT-1988; 88US-00263584.
XX PR 14-NOV-1988; 88US-00271450.
XX PA (CHIR ) CHIRON CORP.
XX PA (CHIR ) CHIRON CORP.
XX PI Houghton M, Choo QL, Kuo G;
XX DR WPI; 1989-159274/22.
XX DR N-PSDB; AAN92106.
XX
XX Purified hepatitis C virus - and associated nucleic acids and
XX polypeptide(s).
XX Claim 13; Fig 47-1-47-8; 139pp; English.
XX
XX It is the sequence encoded in the open reading frame of hepatitis C virus
XX (HCV) cDNA inserts in clones K9-1 through 15e. It is antigenic and could
XX be used in immunoassay reagents and vaccines and to generate antibodies
XX useful in diagnosis and passive immunotherapy for HCV infection/non-A,
XX non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated
XX on 25-MAR-2003 to correct PI field.)
XX
XX Revised record issued on 09-SEP-2004 : Correction to DE line
XX
XX Sequence 2436 AA;
XX
XX Query Match 11.0%; Score 13; DB 1; Length 2436;
XX Best Local Similarity 100.0%; Pred. No. 0.0051;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX Db 1214 GGVLAALAAAYCLS 1226
XX
XX RESULT 222
XX AAP90288
XX ID AAP90288 standard; protein; 2436 AA.
XX AC AAP90288;
XX
XX DT 25-MAR-2003 (revised)
XX DT 19-JUL-2001 (revised)
XX DT 01-NOV-1989 (first entry)
XX
XX DE Peptide encoded by composite hepatitis C cDNA.
XX
XX HCV Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine.
XX
XX Os Pan troglodytes.
XX
XX PN GB2212511-A.
XX
XX DT 26-JUL-1989.
XX
XX PF 18-NOV-1988; 88GB-00027024.
XX
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 26-OCT-1988; 88US-00263584.
XX
XX PA (CHIR ) CHIRON CORP.
XX
XX PI Houghton M, Choo QL, Kuo G;
XX
```

```
XX DR WPI; 1989-215054/30.
XX DR N-PSDB; AAN90336.
XX
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes
XX polypeptide(s) and antibodies for diagnosis, prevention and treatment of
XX infection.
XX
XX PS Disclosure; Fig 47-1 to 47-8; 30pp; English.
XX
XX The sequence is the peptide encoded by the composite hepatitis C virus
XX (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced
XX NANBH, to raise antibodies for immunoassay or treatment, or to produce
XX vaccines. (N.B. This record was resubmitted to correct errors in the
XX sequence.) (Updated on 25-MAR-2003 to correct PR field.)
XX
XX SQ Sequence 2436 AA;
XX
XX Query Match 11.0%; Score 13; DB 1; Length 2436;
XX Best Local Similarity 100.0%; Pred. No. 0.0051;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX Db 1214 GGVLAALAAAYCLS 1226
XX
XX RESULT 223
XX AAR28582
XX ID AAR28582 standard; protein; 2436 AA.
XX AC AAR28582;
XX
XX DT 25-MAR-2003 (revised)
XX DT 22-MAR-1993 (first entry)
XX
XX DE HCV amino acid sequence contg. antibody reactive peptides.
XX
XX KW Hepatitis C virus; detection; peptides.
XX
XX OS Hepatitis C virus.
XX
XX FH Key Location/Qualifiers
XX FT Peptide 828..838
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1236..1275
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1236..1261
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1254..1275
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1270..1295
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1288..1312
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1305..1325
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX FT Peptide 1318..1340
XX FT Peptide /note= "claimed peptide reactive to HCV antibody"
XX
XX PN JP04288097-A.
XX
XX PD 13-OCT-1992.
XX
XX PF 07-NOV-1990; 90JP-00301705.
XX
XX PR 07-NOV-1990; 90JP-00301705.
XX
XX PA (OLYU ) OLYMPUS OPTICAL CO LTD.
XX
XX DR WPI; 1992-387721/47.
XX
XX PT New peptide(s) - are reactive to an antibody against type C hepatitis
```


PD 19-SEP-1990.
 XX 16-MAR-1990; 90EP-00302866.
 XX 17-MAR-1989; 89US-00325338.
 XX 20-APR-1989; 89US-00341334.
 XX 18-MAY-1989; 89US-00355002.
 XX (CHIR) CHIRON CORP.
 XX Houghton M, Choo QL, Kuo G;
 XX WPI; 1990-284418/38.
 XX N-PSDB; AAQ05955.
 XX Hepatitis C virus DNA - used for producing probes, polypeptide(s),
 XX antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.
 XX Disclosure; Fig 16; 83pp; English.
 XX HCV cDNA libraries were constructed using pooled serum from a chimpanzee
 XX with chronic HCV infection. A lambda gt11 library was screened with
 XX probes derived from previously isolated clones. The ORF is derived from
 XX the overlapping clones p14s, CA167b, CA156e, CA84s, CA59a, K9-1, 12f,
 XX 14i, 11b, 7f, 8i, 33c, 40b, 37b, 35, 36, 8i, 32, 33b, 25c, 14c, 8f, 33f,
 XX 33g, 39c, 35f, 19g, 26g and 15e. This polypeptide can be used to design
 XX probes for the detection of HCV nucleic acids, in screening programmes
 XX for antiviral agents and in preparing blood free of HCV. Antisense
 XX polynucleotides can be used to inhibit viral replication. See also
 XX AAQ05956. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX Sequence 2772 AA;
 Query Match 11.0%; Score 13; DB 2; Length 2772;
 Best Local Similarity 100.0%; Pred. No. 0.0056;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAAALAAAYCLS 30
 Db 1550 GGVLAAALAAAYCLS 1562
 RESULT 226
 AAB18540
 ID AAB18540 standard; protein; 2772 AA.
 XX
 XX AAB18540;
 XX
 XX 15-JAN-2001 (first entry)
 XX Protein encoded by a cDNA compiled Hepatitis C virus cDNA clones.
 XX Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;
 XX viral infectivity; viral replication.
 XX Hepatitis C virus.
 XX EP1034785-A2.
 XX
 XX 13-SEP-2000.
 XX 16-MAR-1990; 2000EP-00109602.
 XX 17-MAR-1989; 89US-00325338.
 XX 20-APR-1989; 89US-00341334.
 XX 18-MAY-1989; 89US-00355002.
 XX 16-MAR-1990; 90EP-00302866.
 XX (CHIR) CHIRON CORP.
 XX Houghton M, Choo Q, Kuo G;
 XX WPI; 2000-566891/53.

DR N-PSDB; AAA75296.
 XX Novel composition comprising a hepatitis C virus antisense polynucleotide
 XX which is complementary to or corresponds to a sense strand of the virus
 XX genome, and selectively hybridizes to it.
 XX Example; Fig 16; 75pp; English.
 XX The specification describes a pharmaceutical composition which comprises
 XX a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
 XX characterized by a positive stranded RNA genome which has 40% homology at
 XX the polypeptide level to a HCV polypeptide. The antisense polynucleotide
 XX binds to cellular polynucleotides which enhance and/or are required for
 XX viral infectivity, replicative ability or chronicity. The antisense
 XX polynucleotides may also be designed to bind with high specificity, to be
 XX of increased stability, to be stable and to have low toxicity. The
 XX composition also comprises an agent which causes viral RNA to be
 XX inactive. The composition is used for preventing HCV replication in a
 XX system. The present sequence is encoded by a novel HCV cDNA sequence,
 XX which is used in the course of the invention
 XX Sequence 2772 AA;
 Query Match 11.0%; Score 13; DB 3; Length 2772;
 Best Local Similarity 100.0%; Pred. No. 0.0056;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAAALAAAYCLS 30
 Db 1550 GGVLAAALAAAYCLS 1562
 RESULT 227
 ADN35976
 ID ADN35976 standard; protein; 2772 AA.
 XX
 XX ADN35976;
 XX 17-JUN-2004 (first entry)
 XX HCV cDNA clone #1 protein.
 XX Antiviral; Vaccine; hepatitis C virus infection; HCV infection.
 XX Hepatitis C virus.
 XX EP1394255-A2.
 XX 03-MAR-2004.
 XX 16-MAR-1990; 2003EP-00016585.
 XX 17-MAR-1989; 89US-00325338.
 XX 20-APR-1989; 89US-00341334.
 XX 18-MAY-1989; 89US-00355002.
 XX 16-MAR-1990; 90EP-00302866.
 XX (CHIR) CHIRON CORP.
 XX Houghton M, Choo Q, Kuo G;
 XX WPI; 2004-193149/19.
 XX N-PSDB; ADN35977.
 XX Novel purified hepatitis C virus polypeptide comprising epitope encoded
 XX by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C
 XX virus.
 XX Example 1; Fig 16; 79pp; English.
 XX The present invention relates to hepatitis C virus (HCV) proteins and
 XX cDNA sequences. The sequences are useful in immunoassays for detecting
 XX antibodies directed against HCV antigen; preparing host cells transformed

CC with a recombinant polynucleotide; screening antiviral agents and
 CC determining the effect of antiviral agent in inhibiting viral replication
 CC in cell culture system; and developing vaccine for treating HCV
 CC infection.

XX SQ Sequence 2772 AA;

Query Match 11.0%; Score 13; DB 8; Length 2772;
 Best Local Similarity 100.0%; Pred. No. 0.0056;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 GGVLAALAAAYCLS 30

Db 1550 GGVLAALAAAYCLS 1562

RESULT 228

AAR34009

ID AAR34009 standard; protein; 2816 AA.

XX AC AAR34009;

XX DT 25-MAR-2003 (revised)

XX DT 26-JUL-1993 (first entry)

XX DE HCV-1 polypeptide.

XX Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; HCV;
 KW asymptomatic; chronically infected; epitope; viral isolate; domain;
 KW immunological; cross-reactive; envelope protein; vaccine;
 KW SP53(BVDV)/gp55; hog cholera virus; pestivirus; NS1; flavivirus.

XX Hepatitis C virus.

XX WO9306126-A1.

XX PD 01-APR-1993.

XX PF 11-SEP-1992; 92WO-US007683.

XX PR 13-SEP-1991; 91US-00759575.

XX PA (CHIR) CHIRON CORP.

XX PI Weiner AJ, Houghton M;

XX DR WPI; 1993-117468/14.

XX Immuno-reactive hepatitis C virus polypeptide compans. - contg. at least
 PT 2 sequences from the first variable domain of distinct HCV isolates.

XX PS Disclosure; Fig 9; 106pp; English.

XX This sequence represents the entire hepatitis C virus polypeptide. HCV is
 CC a member of the flavivirus family and appears to encode a basic
 CC polypeptide domain ("C") at the N-terminal of the viral polypeptide,
 CC followed by two glycoprotein domains ("E1", "E2/NS1"), upstream of the
 CC nonstructural genes NS2 through NS5. See also AAQ35134-48, AAR33982-4008
 CC and AAR38088-89. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 2816 AA;

Query Match 11.0%; Score 13; DB 2; Length 2816;
 Best Local Similarity 100.0%; Pred. No. 0.0057;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 GGVLAALAAAYCLS 30

Db 1664 GGVLAALAAAYCLS 1676

RESULT 229

AAR24440

ID AAR24440 standard; protein; 2894 AA.

XX AC AAR24440;

XX DT 25-MAR-2003 (revised)

XX DT 02-DEC-1992 (first entry)

XX DE Composite HCV HC-J1/CDC/CHI protein.

XX KW Hepatitis C virus; peptides; antibodies; ELISA.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 FT Peptide 1..20
 FT Peptide /label= 1

FT Peptide 7..26

FT Peptide /label= 2

FT Peptide 13..32

FT Peptide /label= 3

FT Peptide 37..56

FT Peptide /label= 4

FT Peptide 49..68

FT Peptide /label= 5

FT Peptide 61..80

FT Peptide /label= 6

FT Peptide 73..92

FT Peptide /label= 7

FT Peptide 1688..1707

FT Peptide /label= 8

FT Peptide 1694..1713

FT Peptide /label= 9

FT Peptide 1706..1725

FT Peptide /label= 10

FT Peptide 1712..1731

FT Peptide /label= 11

FT Peptide 1718..1737

FT Peptide /label= 12

FT Peptide 1724..1743

FT Peptide /label= 13

FT Peptide 1730..1749

FT Peptide /label= 14

FT Peptide 2263..2282

FT Peptide /label= 15

FT Peptide 2275..2294

FT Peptide /label= 16

FT Peptide 2287..2306

FT Peptide /label= 17

FT Peptide 2299..2318

FT Peptide /label= 18

FT Peptide 2311..2330

FT Peptide /label= 19

XX EP489968-A1.

XX 17-JUN-1992.

XX 14-DEC-1990; 90EP-00124241.

XX 14-DEC-1990; 90EP-00124241.

XX (INNO-) INNOGENETICS NV.

XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;

XX WPI; 1992-201383/25.

XX New synthetic peptides for detecting antibodies to hepatitis C virus -
 useful in e.g. ELISA assays, and for detection of HCV antigens or as
 immunogens.

XX Disclosure; Fig 1; 32pp; English.

CC RNA viruses frequently exhibit a high rate of spontaneous mutation, thus
CC a virus is considered to be the same of equiv. to HCV if it exhibits a
CC global homology of more than 70 percent with the HCV HC- J1/CDC/CHI
CC composite sequence. The peptide fragments of this DNA sequence indicated
CC in the features table can immunologically mimic proteins encoded by HCV.
CC Additional amino acids or chemical gps. may be added to either end of the
CC peptides for the purpose of creating a linker arm for attachment to a
CC carrier. The peptides can be used for the detection of antibodies
CC specific for HCV. They may be used in the form of kits, opt. with
CC reagents such as staphylococcal protein A, streptococcal protein G,
CC avidin or streptavidin. The peptides may also be used as immunogens for
CC raising antibodies. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 2894 AA;

Query Match 11.0%; Score 13; DB 2; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 1664 GGVLAALAAAYCLS 1676

RESULT 230

AAR70230
ID AAR70230 standard; protein; 2894 AA.

AC AAR70230;

XX 25-MAR-2003 (revised)

DT 07-NOV-1995 (first entry)

XX Composite hepatitis C virus (HC-J1/CDC/CHI).

XX Composite hepatitis C virus; HC-J1/CDC/CHI; HCV; non-A non-B;
XX synthetic antigens; blood screening.

OS Hepatitis C virus.

XX EP644202-A1.

XX 22-MAR-1995.

XX 14-DEC-1990; 94EP-00108611.

XX 14-DEC-1990; 90EP-00124241.

PA (INNO-) INNOGENETICS NV.

XX Deleys RJ, Pollat D, Maertens G, Van Heuverswyn H;

XX WPI; 1995-116946/16.

XX Synthetic antigens for the detection of hepatitis C virus antibodies -
PT comprise portions of the HCV peptide sequence, for use in screening blood
PT and blood products.

PS Disclosure; Fig 1; 51pp; English.

XX AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from
CC which the synthetic HCV antigens described in AAR70210-R70229 were
CC derived. These synthetic antigens can be used to screen blood, or blood
CC products for the presence HCV, they can also be used in various specific
CC assays for the detection of HCV antibodies, and antigens, or as
CC immunogens. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-
CC MAR-2003 to correct PF field.)
XX

SQ Sequence 2894 AA;

Query Match 11.0%; Score 13; DB 2; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 1664 GGVLAALAAAYCLS 1676

RESULT 231
AAR08124
ID AAR08124 standard; protein; 2955 AA.

XX AC AAR08124;

XX 25-MAR-2003 (revised)

DT 23-JAN-1991 (first entry)

XX Hepatitis C virus putative polyprotein.

XX Hepatitis C virus (HCV); antiviral agent.

XX Hepatitis C virus.

OS Key Location/Qualifiers
XX Key Misc-difference 9. .9 /label= K or R
XX Misc-difference 11. .11 /label= N or T
XX Misc-difference 176. .176 /label= I or T
XX Misc-difference 334. .334 /label= M or V
XX Misc-difference 603. .603 /label= I or L
XX Misc-difference 848. .848 /label= Y or N
XX Misc-difference 1114. .1114 /label= P or S
XX Misc-difference 1117. .1117 /label= S or T
XX Misc-difference 1276. .1276 /label= P or L
XX Misc-difference 1454. .1454 /label= C or Y
XX Misc-difference 1471. .1471 /label= T or S
XX Misc-difference 1877. .1877 /label= E or G
XX Misc-difference 1948. .1948 /label= L or H
XX Misc-difference 1949. .1949 /label= S or C
XX Misc-difference 2021. .2021 /label= V or G
XX Misc-difference 2349. .2349 /label= T or S
XX Misc-difference 2385. .2385 /label= Y or F
XX Misc-difference 2386. .2386 /label= S or A
XX Misc-difference 2502. .2502 /label= L or P
XX Misc-difference 2690. .2690 /label= R or G
XX Misc-difference 2921. .2921 /label= R or G

PN EP388232-A.

XX 19-SEP-1990.

XX 16-MAR-1990; 90EP-00302866.

XX 17-MAR-1989; 89US-00325338.

XX 20-APR-1989; 89US-00341334.

XX The specification describes a pharmaceutical composition which comprises
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
CC characterized by a positive stranded RNA genome which has 40% homology at
CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide
CC binds to cellular polynucleotides which enhance and/or are required for
CC viral infectivity, replicative ability or chronicity. The antisense
CC polynucleotides may also be designed to bind with high specificity, to be
CC of increased stability, to be stable and to have low toxicity. The
CC composition also comprises an agent which causes viral RNA to be
CC inactive. The composition is used for preventing HCV replication in a
CC system. The present sequence is encoded by a novel HCV cDNA sequence,
CC which is used in the course of the invention
XX

SQ Sequence 2955 AA;

Query Match 11.0%; Score 13; DB 3; Length 2955;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCIS 30
Db 1664 GGVLAALAAAYCIS 1676
|||||

RESULT 234
ADN35978
ID ADN35978 standard; protein; 2955 AA.
XX
AC ADN35978;
XX
DT 17-JUN-2004 (first entry)
XX
DE HCV cDNA clone #2 protein.
XX
KW Antiviral; Vaccine; hepatitis C virus infection; HCV infection.
XX
OS Hepatitis C virus.
XX
PN EP1394255-A2.
XX
PD 03-MAR-2004.
XX
PF 16-MAR-1990; 2003EP-00016585.
XX
PR 17-MAR-1989; 89US-00325338.
PR 20-APR-1989; 89US-00341334.
PR 18-MAY-1989; 89US-00355002.
PR 16-MAR-1990; 90EP-00302866.
XX
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo Q, Kuo G;
XX
DR WPI; 2004-193149/19.
DR N-PSDB; ADN35979.
XX
PT Novel purified hepatitis C virus polypeptide comprising epitope encoded
PT by hepatitis C virus cDNA, useful as vaccine for treating hepatitis C
PT virus.
XX
PS Example 1; Fig 17; 79pp; English.
XX
CC The present invention relates to hepatitis C virus (HCV) proteins and
CC cDNA sequences. The sequences are useful in immunoassays for detecting
CC antibodies directed against HCV antigen; preparing host cells transformed
CC with a recombinant polynucleotide; screening antiviral agents and
CC determining the effect of antiviral agent in inhibiting viral replication
CC in cell culture system; and developing vaccine for treating HCV
CC infection.
XX
SQ Sequence 2955 AA;

Query Match 11.0%; Score 13; DB 8; Length 2955;
Best Local Similarity 100.0%; Pred. No. 0.0059;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCIS 30
Db 1664 GGVLAALAAAYCIS 1676
|||||

RESULT 235
AAE00449
ID AAE00449 standard; protein; 2984 AA.
XX
AC AAE00449;
XX
DT 19-JUN-2001 (first entry)
XX
DE Hepatitis C virus H77C protein lacking HVR1 from chimpanzee #96A008.
XX
KW Hepatitis C virus; HCV; hypervariable region one; HVR1; vaccine;
KW antiviral; gene therapy; envelope 2 protein; E2; immunisation; mutant;
KW HCV infection; viral replication; passive immunoprophylaxis; mutein.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Region 192..383
FT /label= E1_protein
FT Misc-difference 253
FT /note= "Wild type Lys substituted with Asn"
FT Region 384..746
FT /label= E2_protein
FT /note= "E2 protein lacks hypervariable region 1"
FT Misc-difference 487
FT /note= "Wild type Val substituted with Met"
FT Misc-difference 1116
FT /note= "Wild type Arg substituted with His"
FT Misc-difference 1283
FT /note= "Wild type Tyr substituted with His"
XX
PN WO200121807-A1.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-US025987.
XX
PR 23-SEP-1999; 99US-0155823P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fornas X, Bukh J, Emerson SU, Purcell RH;
XX
DR WPI; 2001-266076/27.
DR N-PSDB; AAD03808.
XX
PT Novel nucleic acid molecules that encode hepatitis C virus envelope 2
PT protein lacking all or part of hypervariable region 1 of envelope
PT protein, useful as vaccine components for treating or preventing HCV
PT infections.
XX
PS Example 5; Page; 80pp; English.
XX
CC The present sequence is Hepatitis C virus (HCV) H77C protein from
CC chimpanzee #96A008 which lacks the hypervariable region one (HVR1) of HCV
CC envelope 2 (E2). The HCV E2 protein lacking HVR1 (H77C(HVR1)) DNA is
CC useful for producing infectious HCV and chimeric HCV viruses which are
CC are useful for identifying cell lines capable of supporting the
CC replication of viruses. The infectious HCV and HVR1- chimeric HCV are
CC used in the production of attenuated or inactivated vaccines which are
CC useful for treating or preventing HCV in a mammal by immunisation. The
CC host cells expressing the H77C(HVR1)DNA is useful as an immunogen to
CC stimulate a protective immune response to HCV. The immunogens are useful
CC for producing protective antibodies to HCV. The antibodies produced are

CC used in passive immunoprophylaxis for treatment of diseases caused by HCV
 CC in animals, especially humans. The H77C(HVR1) DNA is also useful in gene
 CC therapy. Note: The present sequence is not found in the specification but
 CC is derived from Hepatitis C virus envelope protein 2 lacking
 CC hypervariable region 1 (AAE00442) referred as SEQ ID NO: 2 and shown in
 CC figure 1
 CC

SQ Sequence 2984 AA;
 Query Match 11.0%; Score 13; DB 4; Length 2984;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||
 Db 1637 GGVLAALAAAYCLS 1649

RESULT 236
 AAE00447
 ID AAE00447 standard; protein; 2984 AA.
 AC AAE00447;
 DT 19-JUN-2001 (first entry)
 XX Hepatitis C virus H77C protein lacking HVR1 region from chimpanzee 1590.
 XX

DE Hepatitis C virus; HCV; hypervariable region one; HVR1; vaccine;
 KW antiviral; gene therapy; envelope 2 protein; E2; immunisation; mutant;
 KW HCV infection; viral replication; passive immunoprophylaxis; mutain.
 XX

OS Hepatitis C virus.

XX Key Location/Qualifiers
 XX Region 192..383
 FT /label= E1 protein
 FT Region 384..746
 FT /label= E2 protein
 FT /note= "E2 protein lacks hypervariable region 1"
 FT Misc-difference 487
 FT /note= "Wild type Val substituted with Met"
 FT Misc-difference 588
 FT /note= "Wild type Leu substituted with His"
 FT Misc-difference 1116
 FT /note= "Wild type Arg substituted with His"
 FT Misc-difference 1429
 FT /note= "Wild type Thr substituted with Met"
 FT Misc-difference 2848
 FT /note= "Wild type Glu substituted with Asp"

XX WO200121807-A1.

XX 29-MAR-2001.

XX 22-SEP-2000; 2000WO-US025987.

XX 23-SEP-1999; 99US-0155823P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Forns X, Bukh J, Emerson SU, Purcell RH;

XX WPI; 2001-266076/27.

XX N-PSDB; AAD03807.

XX Novel nucleic acid molecules that encode hepatitis C virus envelope 2
 PT protein lacking all or part of hypervariable region 1 of envelope
 PT protein, useful as vaccine components for treating or preventing HCV
 PT infections.

XX Example 4; Page; 80pp; English.

XX

CC The present sequence is Hepatitis C virus (HCV) H77C protein from
 CC chimpanzee 1590 which lacks the hypervariable region one (HVR1) of HCV
 CC envelope 2 (E2). The HCV E2 protein lacking HVR1 (H77C(HVR1)) DNA is
 CC useful for producing infectious HCV and chimeric HCV viruses which are
 CC are useful for identifying cell lines capable of supporting the
 CC replication of viruses. The infectious HCV and HVR1- chimeric HCV are
 CC used in the production of attenuated or inactivated vaccines which are
 CC useful for treating or preventing HCV in a mammal by immunisation. The
 CC host cells expressing the H77C(HVR1)DNA is useful as an immunogen to
 CC stimulate a protective immune response to HCV. The immunogens are useful
 CC for producing protective antibodies to HCV. The antibodies produced are
 CC used in passive immunoprophylaxis for treatment of diseases caused by HCV
 CC in animals, especially humans. The H77C(HVR1) DNA is also useful in gene
 CC therapy. Note: The present sequence is not found in the specification but
 CC is derived from Hepatitis C virus envelope protein 2 lacking
 CC hypervariable region 1 (AAE00442) referred as SEQ ID NO: 2 and shown in
 CC figure 1
 CC

SQ Sequence 2984 AA;

Query Match 11.0%; Score 13; DB 4; Length 2984;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
 |||||
 Db 1637 GGVLAALAAAYCLS 1649

RESULT 237
 AAE00442

ID AAE00442 standard; protein; 2984 AA.

AC AAE00442;

XX 19-JUN-2001 (first entry)

XX Hepatitis C virus envelope 2 protein lacking hypervariable region 1.

DE Hepatitis C virus; HCV; hypervariable region one; HVR1; vaccine;
 KW antiviral; gene therapy; envelope 2 protein; E2; immunisation;
 KW HCV infection; viral replication; passive immunoprophylaxis.

XX Hepatitis C virus.

XX Key Location/Qualifiers

XX Region 192..383

FT /label= E1 protein

FT Region 384..746

FT /label= E2 protein

FT /note= "E2 protein lacks hypervariable region 1"

XX WO200121807-A1.

XX 29-MAR-2001.

XX 22-SEP-2000; 2000WO-US025987.

XX 23-SEP-1999; 99US-0155823P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Forns X, Bukh J, Emerson SU, Purcell RH;

XX WPI; 2001-266076/27.

XX N-PSDB; AAD03778.

XX Novel nucleic acid molecules that encode hepatitis C virus envelope 2
 PT protein lacking all or part of hypervariable region 1 of envelope
 PT protein, useful as vaccine components for treating or preventing HCV
 PT infections.

XX Claim 1; Fig 1G-H; 80pp; English.

XX

XX The present sequence is infectious hepatitis C virus (HCV) of genotype 1a
CC clone [H7C(HVR1)] which lacks the hypervariable region one (HVR1) of HCV
CC envelope 2 [82]. The HCV E2 protein lacking HVR1 DNA is useful for
CC producing infectious HCV and chimeric HCV viruses which are useful
CC for identifying cell lines capable of supporting the replication of
CC viruses. The infectious HCV and HVR1-chimeric HCV are used in the
CC production of attenuated or inactivated vaccines which are useful for
CC treating or preventing HCV in a mammal by immunisation. The host cells
CC expressing the H7C(HVR1) DNA is useful as an immunogen to stimulate a
CC protective immune response to HCV. The immunogens are useful for
CC producing protective antibodies to HCV. The antibodies produced are used
CC in passive immunoprophylaxis for treatment of diseases caused by HCV in
CC animals, especially humans. The H7C(HVR1) DNA is also useful in gene
CC therapy
XX

SQ Sequence 2984 AA;
Query Match 11.0%; Score 13; DB 4; Length 2984;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1637 GGVLAALAAAYCLS 1649
|||||

RESULT 238
ADM24822
ID ADM24822 standard; protein; 3002 AA.
XX
AC ADM24822;
XX
DT 20-MAY-2004 (first entry)
XX
DE Hepatitis C virus (HCV) polyprotein.
XX
KW Hepatotropic; virucide; antiinflammatory; cytostatic; Hepatitis C virus;
KW HCV; HCV infection; DC-SIGN; DC-SIGN related; DC-SIGNR; therapy;
KW liver disease; hepatitis; cirrhosis; hepatocellular carcinoma;
KW polyprotein.
XX
OS Hepatitis C virus.
XX
PN US2003134297-A1.
XX
PD 17-JUL-2003.
XX
PF 26-JUN-2002; 2002US-00184150.
XX
PR 26-JUN-2001; 2001US-0300971P.
XX
PA (OLSO/) OLSON W C.
PA (MADD/) MADDON P J.
XX
PI Olson WC, Maddon PJ;
XX
XX WPI; 2003-829636/77.
DR GENBANK; AF009606.
XX
XX Inhibiting hepatitis C virus infection of susceptible cell or target cell
XX comprises contacting cell with compound to inhibit binding of hepatitis C
XX virus envelope glycoprotein to specified intracellular adhesion proteins.
XX
XX Claim 100; SEQ ID NO 3; 47pp; English.

XX The invention relates to a method for inhibiting Hepatitis C virus (HCV)
CC infection using dendritic cell-specific intracellular adhesion molecule 3
CC -grabbing noninegrin (DC-SIGN) or DC-SIGN related (DC-SIGNR) protein.
CC The method is useful for inhibiting HCV infection of a susceptible cell
CC or a target cell, e.g. a primary cell, a dendritic cell, endometrial cell
CC in liver or placenta cell. It is useful for diagnosing and treating HCV
CC infection. The method is useful for treating or preventing a liver

CC disease, e.g., hepatitis, cirrhosis and hepatocellular carcinoma. The
CC present sequence is HCV polyprotein. This sequence is used to illustrate
CC the method of the invention.
XX

SQ Sequence 3002 AA;
Query Match 11.0%; Score 13; DB 7; Length 3002;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 239
AAR21519
ID AAR21519 standard; protein; 3011 AA.
XX
AC AAR21519;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 22-JUN-1992 (first entry)
XX
DE Compiled HCV sequence.
XX
KW HCV1; serum; gtl1.
XX
OS Hepatitis C virus type 1.
XX
PH Key Location/Qualifiers
FT Misc-difference 9 /label= ARG
FT Misc-difference 11 /label= THR
FT Misc-difference 176 /label= THR
FT Misc-difference 334 /label= VAL
FT Misc-difference 603 /label= ILE
FT Misc-difference 848 /label= (ASN)
FT Misc-difference 1114 /label= SER
FT Misc-difference 1117 /label= THR
FT Misc-difference 1276 /label= LEU
FT Misc-difference 1328 /label= (VAL)
FT Misc-difference 1454 /label= TYR
FT Misc-difference 1471 /label= (SER)
FT Misc-difference 1877 /label= (GLY)
FT Misc-difference 1948 /label= (HIS)
FT Misc-difference 1949 /label= (CYS)
FT Misc-difference 2021 /label= (VAL)
FT Misc-difference 2349 /label= (SER)
FT Misc-difference 2385 /label= (PHE)
FT Misc-difference 2386 /label= (ALA)
FT Misc-difference 2502 /label= (PHE)
FT Misc-difference 2690

CC and V3 regions that contain the greatest amt. of diversity when compared
 CC to known HCV isolates. The polynucleotides encoding these may be used as
 CC primers, probes, or a nucleic acids. (Updated on 25-MAR-2003 to correct
 CC PR field.) (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-
 CC AUG-2003 to correct OS field.)

XX Sequence 3011 AA;

SQ Query Match 11.0%; Score 13; DB 2; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAALAAAYCLS 30
 Db 1664 GGVLAAALAAAYCLS 1676

RESULT 241

AAR31621
 ID AAR31621 standard; protein; 3011 AA.

XX AC AAR31621;

XX DT 25-MAR-2003 (revised)

DT 24-MAY-1993 (first entry)

XX Hepatitis C virus (HCV) polyprotein.

XX Hepatitis; liver disease; HCV1; monoclonal antibody; epitope;
 KW immobilised reagent; immunoassay; diagnosis; detection; treatment;
 KW infection.

XX Hepatitis C virus type 1.

XX Key Location/Qualifiers

FT Domain 1..191
 /label= C domain
 /note= "nucleocapsid protein"

FT Domain 192..383
 /label= E1
 /note= "virion envelope protein"

FT Domain 384..800
 /label= E2/NS1
 /note= "possible envelope"

FT Domain 800..1050
 /label= NS2
 /note= "unknown function"

FT Domain 1050..1650
 /label= NS3
 /note= "putative protease domain"

FT Domain 1651..2100
 /label= NS4
 /note= "unknown function"

FT Domain 2100..3011
 /label= NS5
 /note= "polymerase"

FT WO9300365-A2.

PN 07-JAN-1993.

XX 24-JUN-1992; 92WO-US005388.

XX 24-JUN-1991; 91US-00722489.

XX (CHIR) CHIRON CORP.

XX Chien DY, Rutter W;

XX WPI; 1993-036334/04.

XX Polypeptide(s) comprising truncated hepatitis C virus sequences - for
 PT detection, prevention and treatment of hepatitis C infection.

XX Claim 1; Fig 1; 80pp; English.

XX This sequence represents the polyprotein of the HCV prototype isolate
 CC HCV1. When compared with all known viral sequences, small but significant
 CC co-linear homologies are observed with the non- structural proteins of
 CC the flavivirus family, and with the pestiviruses. The domains shown in
 CC the features table are however, tentatively assigned. The polyprotein,
 CC and epitopes of it are useful for inducing immunological response in a
 CC subject against HCV. The presence of Abs against HCV can be detected
 CC using an immunoassay. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 2; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAALAAAYCLS 30
 Db 1664 GGVLAAALAAAYCLS 1676

RESULT 242

AAR40119
 ID AAR40119 standard; protein; 3011 AA.

XX AC AAR40119;

XX DT 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 27-JAN-1994 (first entry)

XX HCV genomic amino acid sequence isolated from infected chimpanzee CO.
 DE Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV;
 KW human growth hormone; HGH; secretion signal; fusion protein; vaccine.
 XX Hepatitis C virus; Virus.

XX WO9315193-A1.

PN 05-AUG-1993.

XX 29-JAN-1993; 93WO-US000907.

XX 31-JAN-1992; 92US-00830024.

XX (ABBO) ABBOTT LAB.

XX Casey JM, Bode SL, Zeck BJ, Yamaguchi J, Frail DE, Desai SM;
 PI Devare SG;

XX WPI; 1993-258673/32.

XX New plasmid pHCV-162 is a mammalian expression systems for HCV proteins
 PT - useful for diagnosing HCV infection and as vaccines for preventing HCV
 PT infection.

XX Example 1; Page 29-39; 100pp; English.

XX RNA was isolated from the serum of a chimpanzee (designated "CO")
 CC experimentally infected with HCV and cDNA was prepared from it. The cDNA
 CC was PCR amplified using specific primers with sequences based on the
 CC prototype HCV-1 cDNA sequence (GENBANK M62321). Further amplification
 CC using nested primers resulted in 7 adjacent HCV DNA fragments which could
 CC be assembled into a full-length sequence. The DNA sequence was determined
 CC and translated into the genomic amino acid sequence. Comparison of the CO
 CC genomic amino acid sequence with that from HCV-1 showed 98 amino acid
 CC differences. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24
 CC -OCT-2003 to standardise OS field)

XX Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 243

AAR40120
ID AAR40120 standard; protein; 3011 AA.

XX AAR40120;

XX 24-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 27-JAN-1994 (first entry)

XX HCV genomic amino acid sequence isolated from infected human LG.

XX Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV;

KW human growth hormone; HGH; secretion signal; fusion protein; vaccine.

XX Hepatitis C virus; Virus.

XX WO9315193-A1.

XX 05-AUG-1993.

XX 29-JAN-1993; 93WO-US000907.

XX 31-JAN-1992; 92US-00830024.

XX (ABBO) ABBOTT LAB.

XX Casey JM, Bode SL, Zeck BJ, Yamaguchi J, Frail DE, Desai SM;

PI Devare SG;

XX WPI; 1993-258673/32.

XX New plasmid pHCV-162 is a mammalian expression systems for HCV proteins
PT - useful for diagnosing HCV infection and as vaccines for preventing HCV
PT infection.

XX Example 1; Page 39-49; 100pp; English.

XX RNA was isolated from the plasma of a HCV seropositive human (designated
CC "LG") and cDNA was prepared from it. The cDNA was PCR amplified using
CC specific primers with sequences based on the prototype HCV-1 cDNA
CC sequence (GENBANK M62321). Further amplification using nested primers
CC resulted in 7 adjacent HCV DNA fragments which could be assembled into a
CC full-length sequence. The DNA sequence was determined and translated into
CC the genomic amino acid sequence. Comparison of the LG genomic amino acid
CC sequence with that from HCV-1 showed 134 amino acid differences. (Updated
CC on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to
CC standardise OS field)

XX Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 244

AAR66995

ID AAR66995 standard; protein; 3011 AA.

XX AAR66995;
AC 27-AUG-2003 (revised)
DT 01-AUG-1995 (first entry)
XX Hepatitis C virus gene HC-J1/protein.
XX Hepatitis C virus; HCV gene HC-J1/protein; specific antibodies.
XX Hepatitis C virus.
XX JP06284887-A.
XX 11-OCT-1994.
XX 10-DEC-1993; 93JP-00345753.
XX 10-DEC-1992; 92JP-00360705.
XX (IMMO) IMMUNO JAPAN KK.
XX WPI; 1994-362594/45.
XX N-PSDB; AAQ74770.
XX HCV genes and the corresponding proteins - used in the production of anti
PT -HCV antibodies and the detection of HCV infection.
XX Claim 11; Page 18-32; 35pp; Japanese.
XX AAQ74770 encodes AAR66995 the HC-J1/protein, the cDNA can be used in the
CC construction of an expression vector for the transfection of a host
CC cell. The host cell can then be used in the production of proteins and
CC peptides, useful in the preparation of monoclonal and polyclonal HCV-
CC specific antibodies. (Updated on 27-AUG-2003 to correct OS field.)

SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30

Db 1664 GGVLAALAAAYCLS 1676

RESULT 245

AAR79232

ID AAR79232 standard; protein; 3011 AA.

XX AAR79232;

DT 08-DEC-1995 (first entry)

XX HCV sequence.

XX HCV; vaccine.

XX Hepatitis C virus.

XX WO9520664-A2.

XX 03-AUG-1995.

XX 27-JAN-1995; 95WO-US001087.

XX 28-JAN-1994; 94US-00188281.

XX (ABBO) ABBOTT LAB.

XX Watanabe S, Yamaguchi J, Desai SM, Devare SG;

XX

Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 248
AAW34480
ID AAW34480 standard; protein; 3011 AA.
XX AC AAW34480;
XX DT 25-MAR-2003 (revised)
XX DT 16-MAR-1998 (first entry)
XX DE HCV polyprotein.
XX KW PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
XX KW C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;
XX KW NS4.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
FT Misc-difference 366 /note= "can optionally be Arg"
FT Misc-difference 372 /note= "can optionally be Thr"
FT Misc-difference 867 /note= "can optionally be Thr"
FT Misc-difference 1341 /note= "can optionally be Val"
FT Misc-difference 2148 /note= "can optionally be Ile"
FT Misc-difference 2883 /note= "can optionally be Asn"
FT Misc-difference 3681 /note= "can optionally be Ser"
FT Misc-difference 3690 /note= "can optionally be Thr"
FT Misc-difference 4167 /note= "can optionally be Leu"
FT Misc-difference 4323 /note= "can optionally be Val"
FT Misc-difference 4701 /note= "can optionally be Tyr"
FT Misc-difference 4752 /note= "can optionally be Ser"
FT Misc-difference 5970 /note= "can optionally be Gly"
FT Misc-difference 6183 /note= "can optionally be His"
FT Misc-difference 6186 /note= "can optionally be Cys"
FT Misc-difference 6402 /note= "can optionally be Val"
FT Misc-difference 7386 /note= "can optionally be Ser"
FT Misc-difference 7494 /note= "can optionally be Phe"
FT Misc-difference 7497 /note= "can optionally be Ala"
FT Misc-difference 7845 /note= "can optionally be Phe"
FT Misc-difference 8409 /note= "can optionally be Gly"
FT Misc-difference 9102 /note= "can optionally be Gly"
FT Misc-difference 9327 /note= "can optionally be Pro"

PN US5683864-A.
XX PD 04-NOV-1997.
XX PF 07-JUL-1992; 92US-00910760.
XX PR 18-NOV-1987; 87US-00122714.
XX PR 30-DEC-1987; 87US-00139886.
XX PR 26-FEB-1988; 88US-00161072.
XX PR 06-MAY-1988; 88US-00191263.
XX PR 26-OCT-1988; 88US-00263584.
XX PR 14-NOV-1988; 88US-00271450.
XX PR 17-MAR-1989; 89US-00325338.
XX PR 20-APR-1989; 89US-00341334.
XX PR 21-APR-1989; 89US-00353896.
XX PR 18-MAY-1989; 89US-00355002.
XX PR 04-APR-1990; 90US-00504352.
XX PA (CHIR) CHIRON CORP.
XX FI Kuo G, Houghton M, Choo Q;
XX DR WPI; 1997-548976/50.
XX DR N-PSDB; AAT99981.
XX FT Combination of three hepatitis C virus antigens - used for detection of
FT specific antibodies to diagnose infection.
XX PS Disclosure; Col 25-46; 57pp; English.
XX CC This sequence represents the Hepatitis C virus polyprotein. Fragments of
CC the DNA encoding this sequence can be amplified and used in the
CC combination of HCV antigens of the invention. The HCV antigen combination
CC comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC 1-120 of the HCV polyprotein), or its immunologically reactive fragment
CC containing at least 8 aa. It also comprises two additional antigens from
CC two different polyprotein domains, including at least 8 aa from the NS3,
CC NS4, S or NS5 domains of the polyprotein, corresponding, respectively, to
CC aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polyprotein.
CC Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC regions of the HCV polyprotein. These antigen combinations are used
CC diagnostically to detect anti-HCV antibodies, using any standard
CC immunoassay format. These antigen combinations have a broader range of
CC reactivity with antibodies than any antigen individually. (Updated on 25-
CC MAR-2003 to correct PR field.)
XX SQ Sequence 3011 AA;
Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 249
AAW77397
ID AAW77397 standard; protein; 3011 AA.
XX AC AAW77397;
XX DT 17-OCT-2003 (revised)
XX DT 11-JAN-1999 (first entry)
XX DE Hepatitis C virus H77 polyprotein.
XX KW HCV; therapy; diagnosis; vector; gene therapy; vaccine.
XX OS Hepatitis C virus; isolate H77.
XX FN WO9839031-A1.

```
XX PD 11-SEP-1998.
XX PF 26-FEB-1998; 98WO-US004428.
XX PR 04-MAR-1997; 97US-00811566.
XX PA (UNIW ) UNIV WASHINGTON.
XX PI Rice CM, Kolykhalov AA;
XX DR WPI; 1998-520770/44.
XX DR N-PSDB; AAV59361.
XX CC New hepatitis C virus nucleic acid clones - comprising a 5'-terminal
PT conserved sequence, an open reading frame encoding functional components
PT and a 3'-terminal conserved sequence.
XX PS Disclosure; Page 104-115; 209pp; English.
XX CC This is the amino acid sequence of the polyprotein encoded by hepatitis C
CC virus (HCV) isolate H77 (see AAV59361). Its cleavage products form
CC functional components of HCV virus particles and RNA replication
CC machinery. A genetically engineered HCV nucleic acid clone is claimed
CC that comprises at least a functional portion of the HCV H77 nucleic
CC acid sequence. The invention relates to the determination of functional
CC HCV genomic RNA sequences, to construction of infectious HCV DNA clones,
CC and to the use of the clones, or their derivatives, in therapeutic,
CC vaccine and diagnostic applications. The invention is also directed to
CC HCV vectors, e.g. for gene therapy or gene vaccines. The products and
CC methods can also be used for identifying cell lines or animals that are
CC permissive for infection with HCV, for studying HCV infection, isolating
CC functional components of HCV, and for screening for agents capable of
CC modulating HCV replication in vitro and in vivo. (Updated on 17-OCT-2003
CC to standardise OS field)
XX SQ Sequence 3011 AA;
XX Query Match 11.0%; Score 13; DB 2; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAALAAAYCLS 30
XX Db 1664 GGVLAALAAAYCLS 1676
XX RESULT 250
XX AAW77398
XX ID AAW77398 standard; protein; 3011 AA.
XX AC AAW77398;
XX DT 17-OCT-2003 (revised)
XX DT 11-JAN-1999 (first entry)
XX CC Hepatitis C virus-H CMR polyprotein.
XX CC HCV; therapy; diagnosis; vector; gene therapy; vaccine.
XX OS Hepatitis C virus; H-CMR.
XX PN WO9839031-A1.
XX PD 11-SEP-1998.
XX PF 26-FEB-1998; 98WO-US004428.
XX PR 04-MAR-1997; 97US-00811566.
XX PA (UNIW ) UNIV WASHINGTON.
XX PI Rice CM, Kolykhalov AA;
XX PD 11-SEP-1998.
XX PF 26-FEB-1998; 98WO-US004428.
XX PR 04-MAR-1997; 97US-00811566.
XX PA (UNIW ) UNIV WASHINGTON.
XX PI Rice CM, Kolykhalov AA;
XX DR WPI; 1998-520770/44.
XX DR N-PSDB; AAV59378.
XX CC New hepatitis C virus nucleic acid clones - comprising a 5'-terminal
PT conserved sequence, an open reading frame encoding functional components
PT and a 3'-terminal conserved sequence.
XX PS Disclosure; Page 137-148; 209pp; English.
XX CC This is the amino acid sequence of the polyprotein encoded by hepatitis C
CC virus (HCV) isolate H-CMR (see AAV59378). Its cleavage products form
CC functional components of HCV virus particles and RNA replication
CC machinery. A genetically engineered HCV nucleic acid clone is claimed
CC that comprises at least a functional portion of an HCV nucleic acid
CC sequence. The invention relates to the determination of functional HCV
CC genomic RNA sequences, to construction of infectious HCV DNA clones, and
CC to the use of the clones, or their derivatives, in therapeutic, vaccine
CC and diagnostic applications. The invention is also directed to HCV
CC vectors, e.g. for gene therapy or gene vaccines. The products and methods
CC can also be used for identifying cell lines or animals that are
CC permissive for infection with HCV, for studying HCV infection, isolating
CC functional components of HCV, and for screening for agents capable of
CC modulating HCV replication in vitro and in vivo. (Updated on 17-OCT-2003
CC to standardise OS field)
XX SQ Sequence 3011 AA;
XX Query Match 11.0%; Score 13; DB 2; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAALAAAYCLS 30
XX Db 1664 GGVLAALAAAYCLS 1676
XX RESULT 251
XX AAW40038
XX ID AAW40038 standard; protein; 3011 AA.
XX AC AAW40038;
XX DT 26-MAY-1998 (first entry)
XX DE HCV polyprotein.
XX KW Hepatitis C virus C domain; HCV; C antigen; immunological activity;
XX KW NS3 domain; NS4 domain; S domain; NS5 domain.
XX OS Hepatitis C virus.
XX PH Key Location/Qualifiers
XX FT Domain 1..120
XX FT /label= C_domain
XX FT Modified-site 9
XX FT /note= "As given in the specification this amino acid can
FT also be Arg"
XX FT Modified-site 11
XX FT /note= "As given in the specification this amino acid can
FT also be Thr"
XX FT Domain 120..400
XX FT /label= S_domain
XX FT Modified-site 174
XX FT /note= "As given in the specification this amino acid can
FT also be Thr"
XX FT Modified-site 334
XX FT /note= "As given in the specification this amino acid can
FT also be Val"
XX FT Modified-site 603
XX FT /note= "As given in the specification this amino acid can
FT also be Ile"
XX FT Modified-site 847
```

FT FT /note= "As given in the specification this amino acid can
FT FT also be Asn"
FT FT 1050..1640
FT FT /label= NS3_domain
FT FT Modified-site 1114
FT FT /note= "As given in the specification this amino acid can
FT FT also be Ser"
FT FT Modified-site 1217
FT FT /note= "As given in the specification this amino acid can
FT FT also be Thr"
FT FT Modified-site 1276
FT FT /note= "As given in the specification this amino acid can
FT FT also be Leu"
FT FT Modified-site 1328
FT FT /note= "As given in the specification this amino acid can
FT FT also be Val"
FT FT Modified-site 1452
FT FT /note= "As given in the specification this amino acid can
FT FT also be Tyr"
FT FT Modified-site 1472
FT FT /note= "As given in the specification this amino acid can
FT FT also be Ser"
FT FT 1640..2000
FT FT /label= NS4_domain
FT FT Modified-site 1877
FT FT /note= "As given in the specification this amino acid can
FT FT also be Gly"
FT FT Modified-site 1948
FT FT /note= "As given in the specification this amino acid can
FT FT also be His"
FT FT Modified-site 1949
FT FT /note= "As given in the specification this amino acid can
FT FT also be Cys"
FT FT 2000..3011
FT FT /label= NS5_domain
FT FT Modified-site 2021
FT FT /note= "As given in the specification this amino acid can
FT FT also be Val"
FT FT Modified-site 2348
FT FT /note= "As given in the specification this amino acid can
FT FT also be Ser"
FT FT Modified-site 2385
FT FT /note= "As given in the specification this amino acid can
FT FT also be Phe"
FT FT Modified-site 2386
FT FT /note= "As given in the specification this amino acid can
FT FT also be Ala"
FT FT Modified-site 2502
FT FT /note= "As given in the specification this amino acid can
FT FT also be Phe"
FT FT Modified-site 2690
FT FT /note= "As given in the specification this amino acid can
FT FT also be Gly"
FT FT Modified-site 2921
FT FT /note= "As given in the specification this amino acid can
FT FT also be Gly"
FT FT Modified-site 2996
FT FT /note= "As given in the specification this amino acid can
FT FT also be Pro"
FT FT
FT FT US5712087-A.
FT FT
FT FT 27-JAN-1998.
FT FT
FT FT 12-MAY-1995; 95US-00440519.
FT FT
FT FT 04-APR-1990; 90US-00504352.
FT FT 07-JUL-1992; 92US-00910760.
FT FT (CHIR) CHIRON CORP.
FT FT Kuo G, Houghton M, Choo Q;
FT FT

DR WPI; 1998-119973/11.
XX N-PSDB; AAV09989.
XX Immunocassays for hepatitis C virus antibodies - using combinations of
PT antigenic fragments of HCV polyprotein.
XX
XX Disclosure; Fig 1; 59pp; English.
XX
CC This sequence represents the hepatitis C virus (HCV) polyprotein which is
CC used in the construction of novel combinations of HCV antigens that have
CC a broader range of immunological activity than any single HCV antigen. An
CC example of such an antigen given in this specification comprises a first
CC antigen containing at least 8 amino acids of the C domain of the HCV
CC polyprotein and a second antigen comprising at least 8 amino acids of the
CC NS3 domain, the NS4 domain, the S domain or the NS5 domain of the HCV
CC polyprotein in the form of a fusion protein, a physical mixture or bound
CC to a solid matrix. Note: The features given in the specification as
CC represented in the feature table of AAW40038 differ from the positions
CC indicated in Figure 1
XX
SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||||

RESULT 252
AAW98021
ID AAW98021 standard; protein; 3011 AA.
XX
XX AAW98021;
XX
XX 21-JUN-1999 (first entry)
XX
XX Infectious hepatitis C virus genotype 1a/1b chimeric protein.
XX
XX HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
XX assay; antiviral; virucide.
XX
XX Hepatitis C virus.
XX
XX WO9904008-A2.
XX
XX 28-JAN-1999.
XX
XX 16-JUL-1998; 98WO-US014688.
XX
XX 18-JUL-1997; 97US-0053062P.
XX 27-JAN-1998; 98US-00014416.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX
XX WPI; 1999-132252/11.
DR N-PSDB; AAX24833.
XX
XX New isolated hepatitis C virus nucleic acids - used to develop products
XX for the diagnosis, prevention and treatment of HCV infections and for
XX developing screening assays.
XX
XX Claim 7; Fig 16G-H; 126pp; English.
XX
XX This polypeptide is encoded by the genome (see AAX24833) of infectious
XX hepatitis C virus (HCV) chimeric 1a/1b clone pH77C-J4. It comprises the
XX nonstructural region of infectious genotype 1a strain H77 (see AAW98020)
XX and the structural region of infectious genotype 1b strain HC-J4 (see
XX AAW98021). The construction of such chimeric nucleic acid sequences is

CC expected to be of importance in studying the growth and virulence
 CC properties of HCV and in the production of HCV suitable for conferring
 CC protection against multiple genotypes of HCV. The invention also relates
 CC to the introduction of mutations or deletions into infectious nucleic
 CC acid sequences in order to produce an attenuated HCV virus suitable for
 CC vaccine development. Infectious nucleic acid sequences can also be used
 CC to produce attenuated virus via passage in vitro or in vivo of the
 CC viruses produced by transfection of a host cell with the infectious
 CC nucleic acid sequence. Vaccines comprising one or more polypeptides made
 CC from the infectious nucleic acid sequence are used to immunise mammals,
 CC especially humans, against hepatitis C. The nucleic acid sequences can
 CC also be used to induce protective immunity against the virus. The nucleic
 CC acid sequences or their encoded proteases (e.g. NS3 protease) can
 CC additionally be used to develop screening assays to identify antiviral
 CC agents for HCV
 XX
 SQ Sequence 3011 AA;
 Query Match 11.0%; Score 13; DB 2; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 Db 1664 GGVLAALAAAYCLS 1676
 |||||
 RESULT 253
 AAW98020
 ID AAW98020 standard; protein; 3011 AA.
 XX
 AC AAW98020;
 XX
 DT 21-JUN-1999 (first entry)
 XX
 DE Infectious hepatitis C virus genotype 1a strain H77C protein.
 XX
 KW HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
 KW assay; antiviral; virucide.
 XX
 OS Hepatitis C virus.
 XX
 PN WO9904008-A2.
 XX
 PD 28-JAN-1999.
 XX
 PF 16-JUL-1998; 98WO-US014688.
 XX
 PR 18-JUL-1997; 97US-0053062P.
 PR 27-JAN-1998; 98US-00014416.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
 XX
 XX WPI; 1999-132252/11.
 DR N-PSDB; AAX24832.
 XX
 XX New isolated hepatitis C virus nucleic acids - used to develop products
 PT for the diagnosis, prevention and treatment of HCV infections and for
 PT developing screening assays.
 XX
 XX Claim 4; Fig 4G-H; 126pp; English.
 XX
 XX This protein is encoded by the infectious hepatitis C virus (HCV)
 CC genotype 1a strain H77 genome (see AAX24832). H77 was obtained from a
 CC patient in the acute phase of transfusion-associated non-A, non-B
 CC hepatitis. The infectious nucleic acid sequence can be used to produce
 CC chimeric genomes (see AAX24833) consisting of the open reading frames of
 CC infectious nucleic acid sequences of other genotypes (including genotypes
 CC 1-6) and subtypes (such as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV.
 CC The invention also relates to the introduction of mutations or deletions
 CC into infectious nucleic acid sequences in order to produce an attenuated

CC HCV virus suitable for vaccine development. Infectious nucleic acid
 CC sequences can also be used to produce attenuated virus via passage in
 CC vitro or in vivo of the viruses produced by transfection of a host cell
 CC with the infectious nucleic acid sequence. Vaccines comprising one or
 CC more polypeptides made from the infectious nucleic acid sequence are used
 CC to immunise mammals, especially humans, against hepatitis C. The nucleic
 CC acid sequences can also be used to induce protective immunity against the
 CC virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
 CC protease) can additionally be used to develop screening assays to
 CC identify antiviral agents for HCV
 XX
 SQ Sequence 3011 AA;
 Query Match 11.0%; Score 13; DB 2; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 Db 1664 GGVLAALAAAYCLS 1676
 |||||
 RESULT 254
 AAB59173
 ID AAB59173 standard; protein; 3011 AA.
 XX
 AC AAB59173;
 XX
 DT 21-MAR-2001 (first entry)
 XX
 DE Protein encoded by infectious Hepatitis C virus 1a genotype.
 XX
 KW GBV-B; hepatitis C virus; HCV; vaccine.
 XX
 OS Hepatitis C virus.
 XX
 PN WO200075337-A1.
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-US015293.
 XX
 PR 04-JUN-1999; 99US-0137694P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Bukh J, Yanagi M, Emerson SU, Purcell RH;
 XX
 XX WPI; 2001-091214/10.
 XX
 XX New infectious nucleic acids of the GB virus-B clone, useful for
 PT indirectly studying the molecular properties of hepatitis C virus (HCV)
 PT and in developing vaccines and therapeutics for HCV.
 XX
 PS Disclosure; Fig 6; 96pp; English.
 XX
 XX The present invention relates to GB virus-B. The nucleic acid molecules
 CC of the invention are useful for indirectly studying the molecular
 CC properties of hepatitis C virus (HCV). The infectious nucleic acid
 CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
 CC in the development of vaccines and therapeutics for HCV
 XX
 SQ Sequence 3011 AA;
 Query Match 11.0%; Score 13; DB 4; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.006;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCLS 30
 Db 1664 GGVLAALAAAYCLS 1676
 |||||

```
RESULT 255
AAB31169
XX AAB31169 standard; protein; 3011 AA.
XX AC
XX AAB31169;
XX DT
XX 02-APR-2001 (first entry)
XX DE
XX Amino acid sequence of a hepatitis C virus (HCV) clone genotype 1a.
XX KW
XX Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
XX KW HCV; vaccine; viral inhibitor; antiviral.
XX OS
XX Hepatitis C virus.
XX XX
XX Key Location/Qualifiers
XX FT Misc-difference 2532
XX FT /note= "Aa encoded by CCC"
XX XX
XX WO200075352-A2.
XX XX
XX 14-DEC-2000.
XX XX
XX 02-JUN-2000; 2000WO-US015527.
XX XX
XX 04-JUN-1999; 99US-0137817P.
XX XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX XX
XX Nam J, Bukh J, Emerson SU, Purcell RH;
XX XX
XX WPI; 2001-071081/08.
XX DR
XX N-PSDB; AAC86938.
XX XX
XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
XX PT in which the (non-)structural region has been replaced by hepatitis C
XX FT virus (HCV) genome useful for treating or preventing HCV signs and
XX PT symptoms.
XX XX
XX Disclosure; Fig 3G-H; 97pp; English.
XX PS
XX The specification describes a nucleic acid comprising a chimeric virus
XX CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
XX CC the (non-)structural region has been replaced by the (non-)structural
XX CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
XX CC the chimeric virus and the chimeric virus are useful for identifying cell
XX CC lines capable of supporting the replication of these chimeric viruses, in
XX CC screening for neutralizing antibodies to HCV of different genotypes, in
XX CC the production of HCV-BVDV virions, for the development of inactivated or
XX CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
XX CC molecular properties of HCV indirectly in vitro, and in identifying
XX CC inhibitors of viral enzyme activity which would be useful as antiviral
XX CC agents. Formulations or compositions comprising the chimeric viruses may
XX CC be used to treat or prevent the signs and symptoms of HCV. The present
XX CC sequence is encoded by a HCV clone, which is used to construct chimeric
XX CC nucleic acids of the invention
XX SQ
XX Sequence 3011 AA;
XX
XX Query Match 11.0%; Score 13; DB 4; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 1664 GGVLAALAAAYCLS 1676
XX
XX RESULT 256
XX AAU99290
XX ID AAU99290 standard; protein; 3011 AA.
XX XX
XX AC AAU99290;
XX
XX Query Match 11.0%; Score 13; DB 5; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 1664 GGVLAALAAAYCLS 1676
XX
XX RESULT 257
XX AAU84597
XX ID AAU84597 standard; protein; 3011 AA.
XX XX
XX AC AAU84597;
XX XX
XX 08-MAY-2002 (first entry)
XX DT
XX HCV polyprotein 1a (HepC1a) consensus sequence.
XX DE
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX KW viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia.
XX XX
XX Hepatitis C virus.
XX OS
XX WO200190197-A1.
XX PN
XX
```

```

PD 29-NOV-2001.
XX
XX
XX 25-MAY-2001; 2001WO-AU000622.
XX
XX
XX 26-MAY-2000; 2000AU-00007761.
XX
XX (AUSU ) UNIV AUSTRALIAN NAT.
XX
XX Thomson SA, Ramshaw IA;
XX
XX WPI; 2002-147575/19.
XX
XX
XX New synthetic polypeptides having several different segments of at least
XX one parent polypeptide linked together differently compared to the
XX linkage in the parent polypeptide, for inducing immune response against a
XX pathogen or cancer.
XX
XX Example 2; Fig 26; 364pp; English.
XX
XX The invention relates to a new synthetic polypeptide (I) comprising
XX several different segments of at least one parent polypeptide linked
XX together in a different relationship relative to their linkage in the
XX parent polypeptide to impede, abrogate or otherwise alter at least one
XX function associated with the parent polypeptide and for inducing an
XX immune response against a pathogen or cancer. Also included are a
XX synthetic polynucleotide encoding and a computer system for designing the
XX synthetic polypeptides. The synthetic polypeptides and polynucleotides
XX are referred to as a Savine. The synthetic polypeptide is useful for
XX modulating immune responses preferably directed against a pathogen or a
XX cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
XX and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
XX esophagus, brain, testicle, uterus), as potentiating agents.
XX Compositions comprising the polypeptide may be used in the treatment or
XX prophylaxis against viral (such as infections caused by HIV (human
XX immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
XX virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
XX (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
XX Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
XX (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
XX Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
XX a consensus sequence for a parent protein used to design a savine of the
XX invention
XX
XX Sequence 3011 AA;
XX
XX Query Match 11.0%; Score 13; DB 5; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 1664 GGVLAALAAAYCLS 1676
XX
XX RESULT 258
XX AAE22049
XX ID AAE22049 standard; protein; 3011 AA.
XX
XX AC AAE22049;
XX
XX DT 16-JUL-2002 (first entry)
XX
XX Hepatitis C virus (HCV) polypeptide.
XX
XX Hepatitis C virus; HCV; antigen; C domain; polypeptide; NS3 domain;
XX NS4 domain; S domain; NS5 domain.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX Domain 1..122
XX /label= C domain
XX 199..328

```

```

/label= S domain
1192..1931
/label= "C200 polypeptide"
1192..1457
/label= "NS3 domain antigen"
1569..1931
/label= "NS4 antigen"
2054..2464
/label= "NS5 antigen"

US6312889-B1.
XX
XX 06-NOV-2001.
XX
XX 12-MAY-1995; 95US-00440549.
XX
XX 04-APR-1990; 90US-00504352.
XX
XX 07-JUL-1992; 92US-00910760.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo Q, Kuo G;
XX
XX WPI; 2002-040268/05.
XX
XX N-PSDB; AAD35043.
XX
XX Combination of hepatitis C viral (HCV) antigens, useful in improved
XX immunoassay for detecting HCV antibodies.
XX
XX Example 1; Col 45-60; 58pp; English.
XX
XX The invention relates to combination of hepatitis C viral (HCV) antigens
XX that have a broader range of immunological reactivity than any single HCV
XX antigen. The combinations consist of an antigen from the C domain of the
XX HCV polypeptide, and at least one additional HCV antigen from either the
XX NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
XX the form of fusion protein, a simple physical mixture, or the individual
XX antigens commonly bound to a solid matrix. The combinations of antigens
XX provides broad range immunoassays for anti-HCV antibodies. The invention
XX therefore provides a method for detecting antibodies to HCV in a mammal
XX suspected of containing such antibodies. The present sequence is HCV
XX polypeptide. Note: This sequence SEQ ID NO:10 is stated to be similar to
XX the sequence shown in Fig 1 (AAE22052) of the specification. However
XX these sequences differ
XX
XX Sequence 3011 AA;
XX
XX Query Match 11.0%; Score 13; DB 5; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 1664 GGVLAALAAAYCLS 1676
XX
XX RESULT 259
XX AAE22052
XX ID AAE22052 standard; protein; 3011 AA.
XX
XX AC AAE22052;
XX
XX DT 16-JUL-2002 (first entry)
XX
XX Hepatitis C virus (HCV) polypeptide, alternative version.
XX
XX Hepatitis C virus; HCV; antigen; C domain; polypeptide; NS3 domain;
XX NS4 domain; S domain; NS5 domain.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX Domain 1..122

```



```
FT /label= C_domain
FT Misc-difference 9
FT /label= Lys, Arg
FT /note= "Encoded by AAA"
FT Misc-difference 11
FT /label= Thr, Asn
FT /note= "Encoded by AAC"
FT Misc-difference 176
FT /label= Thr, Ile
FT /note= "Encoded by ATC"
FT Domain
FT 199..328
FT /label= S_domain
FT Misc-difference 334
FT /label= Val, Met
FT /note= "Encoded by ATG"
FT Misc-difference 603
FT /label= Ile, Leu
FT /note= "Encoded by ATC"
FT Misc-difference 848
FT /label= Asn, Tyr
FT /note= "Encoded by TAT"
FT Misc-difference 1114
FT /label= Ser, Pro
FT /note= "Encoded by CCG"
FT Misc-difference 1117
FT /label= Thr, Ser
FT /note= "Encoded by AGC"
FT Region
FT 1192..1931
FT /note= "c200 polypeptide"
FT Region
FT 1192..1457
FT /note= "NS3 domain antigen"
FT Misc-difference 1276
FT /label= Leu, Pro
FT /note= "Encoded by CCT"
FT Misc-difference 1328
FT /label= Val, Gly
FT /note= "Encoded by GGC"
FT Misc-difference 1454
FT /label= Tyr, Cys
FT /note= "Encoded by TGC"
FT Misc-difference 1471
FT /label= Ser, Thr
FT /note= "Encoded by ACC"
FT Region
FT 1569..1931
FT /note= "NS4 antigen"
FT Misc-difference 1877
FT /label= Gly, Glu
FT /note= "Encoded by GAG"
FT Misc-difference 1948
FT /label= His, Leu
FT /note= "Encoded by CTC"
FT Misc-difference 1949
FT /label= Cys, Ser
FT /note= "Encoded by AGC"
FT Misc-difference 2021
FT /label= Val, Gly
FT /note= "Encoded by GTG"
FT Region
FT 2054..2464
FT /note= "NS5 antigen"
FT Misc-difference 2349
FT /label= Ser, Thr
FT /note= "Encoded by ACC"
FT Misc-difference 2385
FT /label= Phe, Tyr
FT /note= "Encoded by TAT"
FT Misc-difference 2386
FT /label= Ala, Ser
FT /note= "Encoded by TCC"
FT Misc-difference 2502
FT /label= Phe, Leu
FT /note= "Encoded by TTG"
FT Misc-difference 2690
FT /label= Gly, Arg

/label= "Encoded by AGG"
FT Misc-difference 2921
FT /label= Gly, Arg
FT /note= "Encoded by AGA"
FT Misc-difference 2996
FT /label= Pro, Leu
FT /note= "Encoded by CTA"
FT
FT US6312889-B1.
FT
FT 06-NOV-2001.
FT
FT 12-MAY-1995; 95US-00440549.
FT
FT 04-APR-1990; 90US-00504352.
FT 07-JUL-1992; 92US-00910760.
FT
FT (CHIR ) CHIRON CORP.
FT
FT Houghton M, Choo Q, Kuo G;
FT
FT WPI; 2002-040268/05.
FT N-PSDB; AAD35043.
FT
FT Combination of hepatitis C viral (HCV) antigens, useful in improved
FT immunoassay for detecting HCV antibodies.
FT
FT Example 1; Fig 1; 59pp; English.
FT
FT The invention relates to combination of hepatitis C viral (HCV) antigens
FT that have a broader range of immunological reactivity than any single HCV
FT antigen. The combinations consist of an antigen from the C domain of the
FT HCV polypeptide, and at least one additional HCV antigen from either the
FT NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
FT the form of fusion protein, a simple physical mixture, or the individual
FT antigens commonly bound to a solid matrix. The combinations of antigens
FT therefore provides a method for detecting antibodies to HCV in a mammal
FT suspected of containing such antibodies. The present sequence is HCV
FT polypeptide, alternative version. Note: This sequence SEQ.ID.NO:10 is
FT stated to be similar to the sequence shown in Column 45-60 (AAE22049) of
FT the specification. However these sequences differ
FT
FT Sequence 3011 AA;
FT
FT Query Match 11.0%; Score 13; DB 5; Length 3011;
FT Best Local Similarity 100.0%; Pred. No. 0.006;
FT Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
FT Qy 18 GGVLAALAAVCLS 30
FT |||||
FT Db 1664 GGVLAALAAVCLS 1676
FT
FT RESULT 260
FT AAU79221
FT ID AAU79221 standard; protein; 3011 AA.
FT AC AAU79221;
FT XX
FT 29-AUG-2003 (revised)
FT 15-JUL-2002 (first entry)
FT XX
FT Hepatitis C Virus (HCV) E2 protein.
FT XX
FT Hepatitis C Virus E2; HCV E2; virucide; hepatotropic; IL-2;
FT antiinflammatory; HCV infection; interleukin-2; gamma-interferon;
FT granulocyte macrophage-colony stimulating factor; GM-CSF;
FT delta-delta E2 genotype 1a.
FT XX
FT Hepatitis C virus; Virus.
FT OS
FT WO200222155-A1.
FT FN
```

```
XX PD 21-MAR-2002.
XX PF 13-SEP-2001; 2001WO-US028767.
XX PR 13-SEP-2000; 2000US-0230927P.
XX PA (HAWA-) HAWAII BIOTECHNOLOGY GROUP INC.
XX PI Nakano ET, Clements DB, Humphreys T;
XX PS WPI; 2002-383102/41.
XX DR
XX CC The invention relates to a secreted polypeptide comprising hepatitis C
XX CC virus (HCV) E2 polypeptide lacking all or a portion of its membrane
XX CC spanning domain so that the E2 polypeptide is capable of secretion into
XX CC growth medium when expressed recombinantly in a host cell. The
XX CC polypeptide may also lack a portion of its C-terminus. The HCV E2
XX CC secreted polypeptide is useful for producing anti-HCV antibodies. A
XX CC purified immunogenic polypeptide comprising HCV E2 is useful for treating
XX CC HCV infection and for providing immune protection against HCV infection
XX CC by administering it to a subject having or at risk of having HCV
XX CC infection or in need of protection. The method further comprises
XX CC administering an immunomodulatory agent such as interleukin-2 (IL-2),
XX CC granulocyte macrophage-colony stimulating factor (GM-CSF) or gamma-
XX CC interferon. The polypeptide is useful as a vaccine, and with other HCV
XX CC proteins to form a multi-component HCV vaccine for prophylactic or
XX CC therapeutic treatment of HCV infection. This sequence represents an HCV
XX CC E2 protein. (Updated on 29-AUG-2003 to standardise OS field)
XX SQ Sequence 3011 AA;
Query Match 11.0%; Score 13; DB 5; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCLS 30
Db 1664 GGVLAAALAAAYCLS 1676
RESULT 261
AAE19888
ID AAE19888 standard; protein; 3011 AA.
XX AC AAE19888;
XX DT 18-JUN-2002 (first entry)
XX DE Hepatitis C virus (HCV) protein.
XX DE Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX DE cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX OS Hepatitis C virus.
XX OS WO200213855-A2.
XX FN
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001808.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M, Hultgren C;
XX DR WPI; 2002-241837/29.
XX PR N-PSDB; AAD31764.
XX PT Vaccine compositions for treating and preventing disease, preferably
XX PT hepatitis C virus infection, comprises ribavirin and antigen that has
XX PS epitope present in hepatitis C virus.
XX CC Claim 4; Page 65-72; 120pp; English.
XX CC The invention relates to a composition comprising ribavirin and an
XX CC antigen preferably non structural 3 protein (NS3)/4A fragment of
XX CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX CC sequence. The composition is useful for enhancing an immune response to a
XX CC hepatitis C antigen in humans, domestic, sport or pet species and as
XX CC vaccines for treating and preventing HCV infections. The composition is
XX CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX CC The present sequence is HCV protein
XX SQ Sequence 3011 AA;
Query Match 11.0%; Score 13; DB 5; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCLS 30
Db 1664 GGVLAAALAAAYCLS 1676
RESULT 262
AAO26784
ID AAO26784 standard; protein; 3011 AA.
XX AC AAO26784;
XX DT 03-APR-2003 (first entry)
XX DE Protein derived from Hepatitis C virus DNA, SEQ ID No 2.
XX DE Virucide; hepatotropic; antiinflammatory; prenylation; hepatitis C virus;
XX DE HCV; NS5A protein; anti-HCV agent; HCV infection.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1168 /label= His
XX FT /note= "Residue encoded by CCG"
XX FT Misc-difference 1169 /label= Ala
XX FT /note= "Residue encoded by CTG"
XX FT Misc-difference 1170 /label= Val
XX FT /note= "Residue encoded by TTG"
XX FT Misc-difference 1171 /label= Gly
XX FT /note= "Residue encoded by TGC"
XX FT Misc-difference 1172 /label= Leu
XX FT /note= "Residue encoded by CCC"
XX FT Misc-difference 1173 /label= Phe
XX FT /note= "Residue encoded by GCG"
XX FT Misc-difference 1174 /label= Arg
XX FT /note= "Residue encoded by GGA"
XX FT Misc-difference 1175 /label= Ala
XX FT /note= "Residue encoded by CAC"
XX FT Misc-difference 1178
```



```
XX PI Rice CM, Kolykhalov AA;
XX WPI; 2003-466160/44.
XX N-PSDB; ACA62483.
XX
XX Novel genetically engineered hepatitis C virus, for identifying animal
XX permissive for HCV infection, has positive sense nucleic acid, 5'- and 3'
XX -non-translated regions, open reading frame encoding HCV polyprotein.
XX
XX Example 4; Page 62-70; 120pp; English.
XX
XX The invention relates to a genetically engineered hepatitis C virus (HCV)
XX nucleic acid clone. The genetically engineered HCV is useful for
XX identifying a cell line or animal that is permissive for infection with
XX HCV. The genetically engineered HCV is useful for selecting for HCV with
XX adaptive mutations that permit higher levels of HCV replication in a
XX permissive cell line. The genetically engineered HCV or HCV DNA or RNA is
XX useful for infecting an animal with HCV and for propagating HCV in vitro.
XX The genetically engineered HCV or HCV DNA or RNA is useful for
XX transducing an animal susceptible to HCV infection with a heterologous
XX gene. A host cell line, a non-human animal or an in vitro cell line
XX infected with the HCV genomic RNA of the invention is useful for
XX producing HCV virus particles and for screening for agents capable of
XX modulating HCV replication. The genetically engineered HCV is useful for
XX producing HCV replicon proteins. A HCV virus particle comprising a
XX replication-competent or replication-defective HCV genome RNA is useful
XX for producing antibodies to HCV. A HCV virus particle comprising a
XX replication-competent or replication-defective HCV genome RNA is useful
XX in an in vitro method for detecting antibodies to HCV in a biological
XX sample from a subject. The genetically engineered HCV is useful for
XX therapeutic, vaccine and diagnostic applications. A plasmid encoding the
XX genetically engineered HCV is useful for gene therapy and in gene
XX vaccines. The genetically engineered HCV is useful for expression of
XX functional HCV proteins and infection in vivo and in vitro for
XX development of antiviral therapeutics and diagnostics. A non-human animal
XX or an in vitro cell line infected with the HCV genomic RNA of the
XX invention is useful for studying the natural history of HCV infection,
XX for isolating components of HCV, and for sensitive, fast diagnostic
XX applications. The present sequence represents the amino acid sequence of
XX HCV-H
XX
XX SQ Sequence 3011 AA;
XX
XX Query Match 11.0%; Score 13; DB 6; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 1664 GGVLAALAAAYCLS 1676
XX
XX RESULT 265
XX ABW00339
XX ID ABW00339 standard; protein; 3011 AA.
XX AC ABW00339;
XX XX
XX DT 15-JAN-2004 (first entry)
XX XX
XX DE Hepatitis C virus protein.
XX XX
XX KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX KW virucide.
XX XX
XX OS Hepatitis C virus.
XX XX
XX FH Key Location/Qualifiers
XX Domain 1..182
XX FT /note= "Domain"
XX FT 183..379
XX FT /note= "Domain"
XX FT
XX
XX FT 380..729
XX /note= "Domain"
XX Domain
XX 730..1044
XX /note= "Domain"
XX Domain
XX 1045..1657
XX /note= "Domain"
XX Domain
XX 1658..1711
XX /note= "Domain"
XX Domain
XX 1712..1971
XX /note= "Domain"
XX Domain
XX 1972..3011
XX /note= "Domain"
XX
XX US2002136740-A1.
XX
XX 26-SEP-2002.
XX
XX 15-AUG-2001; 2001US-00929955.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX N-PSDB; AAD60865.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Claim 4; Page 28-35; Opp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
XX antigen, where the antigen is derived from a hepatitis virus. The vaccine
XX is useful in enhancing the immune response to a hepatitis C antigen where
XX the composition is delivered to an animal identified as requiring an
XX enhanced immune response. The vaccine is useful in the treatment and
XX prevention of hepatitis C infection. The present sequence is Hepatitis C
XX virus protein
XX
XX SQ Sequence 3011 AA;
XX
XX Query Match 11.0%; Score 13; DB 7; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 0.006;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCLS 30
XX |||||
XX Db 1664 GGVLAALAAAYCLS 1676
XX
XX RESULT 266
XX ADH79949
XX ID ADH79949 standard; protein; 3011 AA.
XX AC ADH79949;
XX XX
XX DT 22-APR-2004 (first entry)
XX XX
XX DE HCV envelope glycoprotein.
XX XX
XX KW HCV envelope glycoprotein; DC-SIGN; HCV infection; liver disease;
XX KW hepatocellular carcinoma; hepatitis; cirrhosis; DC-SIGNR;
XX KW HCV envelope glycoprotein; cytostatic; antiinflammatory; hepatotropic;
XX KW virucide.
XX XX
XX OS Hepatitis C virus.
XX XX
XX XX US2003232745-A1.
XX PN
```

XX PD 18-DEC-2003.
XX XX
XX PF 24-DEC-2002; 2002US-00328997.
XX XX
XX PR 26-JUN-2001; 2001US-0300971P.
XX PR 26-JUN-2002; 2002US-00184150.
XX XX
XX PA (OLSO/) OLSON W C.
XX PA (MADD/) MADDON P J.
XX PA (GARD/) GARDNER J P.
XX PI Olson WC, Maddon PJ, Gardner JP;
XX XX
XX DR WPI; 2004-061306/06.
XX XX
XX PT Use of DC-SIGN and DC-SIGNR proteins for inhibiting, preventing or
XX PT treating HCV infection and liver disease e.g. hepatitis or cirrhosis or
XX PT hepatocellular carcinoma.
XX XX
XX PS Claim 100; SEQ ID NO 3; 55pp; English.
XX XX
XX CC The invention relates to the use of DC-SIGN and DC-SIGNR proteins for
XX CC inhibiting, treating or preventing HCV infection and liver diseases such
XX CC as hepatocellular carcinoma, hepatitis or cirrhosis. The DC-SIGN and DC-
XX CC SIGNR protein are useful for inhibiting HCV infection, treating HCV
XX CC infection or treating or preventing liver disease e.g. hepatitis,
XX CC cirrhosis or hepatocellular carcinoma. Antibodies to the polypeptides,
XX CC the polypeptides or non-peptidyl agents can be used for inhibiting
XX CC binding of a DC-SIGNR or DC-SIGN protein to an HCV envelope glycoprotein.
XX CC This sequence represents the E2 HCV envelope glycoprotein of the
XX CC invention.
XX XX
XX SQ Sequence 3011 AA;
XX XX

Query Match 11.0%; Score 13; DB 8; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
XXXXXXXXXXXX

RESULT 267
ADJ56744
ID ADJ56744 standard; protein; 3011 AA.
XX AC ADJ56744;
XX XX
XX DT 06-MAY-2004 (first entry)
XX XX
XX DE Hepatitis C virus protein SeqID 14.
XX XX
XX KW GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
XX KW vaccine; virucidal; antiinflammatory.
XX OS Hepatitis C virus.
XX OS
XX XX WO2004005498-A1.
XX XX
XX PD 15-JAN-2004.
XX XX
XX XX 02-JUL-2003; 2003WO-US021002.
XX XX
XX PR 03-JUL-2002; 2002US-00189359.
XX XX
XX PA (TEXA) UNIV TEXAS SYSTEM.
XX PA (INSP) INST PASTEUR.
XX XX
XX PI Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX XX
XX DR WPI; 2004-091362/09.

DR N-ESDB; ADJ56743.
XX XX
XX PT New chimeric GBV-B polynucleotide, useful as a model for hepatitis C
XX PT virus, for identifying compounds active against a viral infection, or for
XX PT developing hepatitis C virus preventive and therapeutic treatments.
XX PS Example 23; SEQ ID NO 14; 108pp; English.
XX XX
XX CC This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
XX CC polynucleotides. Specifically, it refers to using the hepatotropic
XX CC flavivirus GBV-B that has a unique phylogenetic relationship to the human
XX CC hepatitis C virus (HCV) and can serve as a surrogate virus in drug
XX CC discovery efforts related to antiviral drug development. The present
XX CC invention describes the construction of an infectious molecular clone
XX CC using the newly determined 3' terminal sequence of GBV-B. Furthermore,
XX CC the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
XX CC envelope proteins such that they can have utility as a vaccine immunogen
XX CC for hepatitis C. In addition, they can be used for screening compounds
XX CC active against viral infection, as well as for developing HCV
XX CC preventative and therapeutic treatments. Accordingly, these compositions
XX CC exhibit virucidal, antiinflammatory and hepatotropic activities. This
XX CC polypeptide sequence is the hepatitis C virus protein sequence of the
XX CC invention.
XX XX
XX SQ Sequence 3011 AA;
XX XX

Query Match 11.0%; Score 13; DB 8; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
XXXXXXXXXXXX

RESULT 268
ADL23107
ID ADL23107 standard; protein; 3011 AA.
XX AC ADL23107;
XX XX
XX DT 20-MAY-2004 (first entry)
XX XX
XX DE Hepatitis C virus protein sequence SeqID 2.
XX KW hepatitis C virus; HCV; viral envelope; E2 glycoprotein;
XX KW low density lipoprotein; LDL; HCV infection; hypercholesterolaemia;
XX KW hyperlipidaemia; coronary heart disease; hepatotropic; virucidal;
XX KW plasma lipoprotein.
XX OS Hepatitis C virus.
XX OS
XX PN WO2004003141-A2.
XX XX
XX PD 08-JAN-2004.
XX XX
XX PF 24-JUN-2003; 2003WO-US019834.
XX XX
XX PR 28-JUN-2002; 2002US-0392158P.
XX PR 27-MAY-2003; 2003US-00445724.
XX XX
XX PA (IOWA) UNIV IOWA RES FOUND.
XX XX
XX PI Stapleton JT, Wuenschmann S;
XX XX
XX DR WPI; 2004-083029/08.
XX DR N-ESDB; ADL23106.
XX XX
XX PT Use of Hepatitis C Virus E2 glycoprotein in reducing low density
XX PT lipoprotein levels in a subject or in treating or preventing hepatitis C
XX PT virus infection.
XX XX
XX PS Disclosure; SEQ ID NO 2; 135pp; English.

XX This invention relates to a novel method for identifying inhibitors of
CC the hepatitis C virus (HCV), and also the use of an HCV viral envelope
CC protein identified as E2 glycoprotein in reducing low density lipoprotein
CC (LDL) levels. Specifically, it refers to the knowledge that HCV E2
CC glycoprotein binds to the lipid moiety of human lipoproteins and uses the
CC natural LDL receptor to bind to cells, thus identifies a novel route by
CC which HCV gains entry to the cell providing a novel therapeutic target.
CC The present invention describes screening methods to identify inhibitors
CC of HCV infection that comprises administering a candidate substance with the
CC E2 glycoprotein and a plasma lipoprotein in order to determine a
CC reduction in binding in the presence of an appropriate inhibitor. As
CC such, these compositions can be used to reduce hypercholesterolaemia and
CC in turn treat hyperlipidaemia and associated conditions including
CC coronary heart disease. Accordingly, they exhibit hepatotropic and
CC virucidal activities. This polypeptide sequence is the HCV protein
CC sequence of the invention.

XX
SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 8; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
| | | | | | | | | |
Db 1664 GGVLAAALAAAYCLS 1676

RESULT 269
ADJ64256

ID ADJ64256 standard; protein; 3011 AA.

AC ADJ64256;

DT 20-MAY-2004 (first entry)

XX Hepatitis C virus polypeptide.

DE antiinflammatory; hepatotropic; virucide; GB virus B; GBV-B;

XX hepatitis C virus; HCV; 3' terminal; hepatitis C virus; HCV.

OS Hepatitis C virus.

XX US2004039187-A1.

XX 26-FEB-2004.

XX 03-JUL-2002; 2002US-00189359.

XX 04-JUN-1999; 99US-0137665P.

PR 05-JUN-2000; 2000US-00587653.

XX (TEXA) UNIV TEXAS SYSTEM.

PA (INSP) INST PASTEUR.

XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;

PI WPI; 2004-203294/19.

XX New GB virus B and/or hepatitis C virus (HCV) sequences, useful in
PT diagnosing and in treating HCV and in investigating the mechanisms for
PT the different biological properties of the viruses.

XX Disclosure; SEQ ID NO 14; 58pp; English.

XX The invention describes a new isolated polynucleotide (I) encoding a 3'
CC sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
CC GBV-B genome, where at least part, but not all of a 5' nontranslated
CC region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
CC (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
CC fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
CC or chimaeras are useful diagnosing or treating hepatitis C virus (HCV)

CC and in investigating the mechanisms for the different biological
CC properties of the viruses. This is the amino acid sequence of a hepatitis
CC C virus and can be used in the creation of a chimeric virus comprising
CC domain III of the 5'NTR of HCV (IRES) within a genetic background of GBV-
CC B.

XX Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 8; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
| | | | | | | | | |
Db 1664 GGVLAAALAAAYCLS 1676

RESULT 270
ADL72983

ID ADL72983 standard; protein; 3011 AA.

AC ADL72983;

DT 17-JUN-2004 (first entry)

XX Hepatitis C virus E1E2 proteins.

DE HCV; hepatitisvirus; vaccine; hepatitis; infection; plasmid; ds; gene.

OS Hepatitis C virus.

XX Key Location/Qualifiers

FT Region 132

FT Peptide /note= "delta C"

FT Protein 171..191

FT Protein 192..383

FT Protein /note= "E1"

FT Protein 384..746

FT Protein /note= "E2"

FT Protein 384..410

FT Protein /note= "HRV1"

FT Protein 747..809

FT Protein /note= "p"

XX WO2004024904-A2.

XX 25-MAR-2004.

XX 12-SEP-2003; 2003WO-IB003882.

XX 13-SEP-2002; 2002EP-00292254.

PR 03-MAR-2003; 2003EP-00290505.

XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.

XX Bartosch B, Cosset F;

XX WPI; 2004-270036/25.

XX Producing infectious hepatitisvirus-like particles ex vivo, comprising
PT transfecting host cells with nucleic acid sequences and maintaining the
PT transfecting cells in culture for sufficient time to allow expression of
PT the cDNAs.

XX Example 1; Page 107-128; 128pp; English.

XX The present invention relates to a method of producing infectious
CC hepatitisvirus-like particles ex vivo, comprising transfecting host cells
CC with nucleic acid sequences and maintaining the transfecting cells in
CC culture for sufficient time to allow expression of the cDNAs to produce
CC structural proteins from hepatitisvirus and retrovirus. The method is useful
CC for producing infectious hepatitisvirus-like particles. The nucleic acid

CC sequences are useful for the preparation of a medicament useful as a
CC vaccine against hepatitis. The hepatitis-like particle that carries a
CC transgene of interest, for the preparation of a medicament for the
CC prevention or treatment of a disease in a patient, where the hepatitis-
CC like particle allows the transfer of the transgene of interest into a
CC cell of the patient, and encodes a product that has a prophylactic or
CC therapeutic effect against the disease. The present sequence contains a
CC number of HCV proteins shown in the exemplification of the invention.

XX SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 8; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0;

Qy 18 GGVLAAALAAAYCLSL 30
Db 1664 GGVLAAALAAAYCLSL 1676

RESULT 271

ADR29357
ID ADR29357 standard; protein; 3011 AA.

AC ADR29357;

DT 18-NOV-2004 (first entry)

DE Hepatitis C virus polypeptide precursor protein SEQ ID NO:1.

KW anti-hepatitis C virus; anti-HCV; immune response; anti-HIV; virucide;
KW antiinflammatory; hepatotropic; immunostimulant; vaccine; immunotherapy;
KW anti-HCV vaccine; AIDS; hepatitis C virus; HCV; polypeptide precursor.

OS Hepatitis C virus.

FH Key Location/Qualifiers
FT Peptide 2594..2602

FT /note= "specifically claimed peptide in claims 4 and 5"

FN WO2004071414-A2.

PD 26-AUG-2004.

PF 03-FEB-2004; 2004WO-US003044.

PR 05-FEB-2003; 2003US-0445438P.

PR 25-MAR-2003; 2003US-0457870P.

XX (GENZ) GENZYME CORP.

PA (GENO) GEN HOSPITAL CORP DBA MASSACHUSETTS GEN.

PI Nicolette CA, Walker BD;

XX WPI; 2004-625761/60.

DR N-PSDB; ADR29358.

DR GENBANK; M62321.

XX New anti- hepatitis C virus (HCV) (AL9) compound, useful as component of
PT anti-HCV vaccines and to expand immune effector cell specific for cells
PT expressing HCV epitope.

XX Claim 4; SEQ ID NO 1; 69pp; English.

XX The present invention describes anti-hepatitis C virus (HCV) (AL9)
CC compounds (I) having the amino acid sequences of SEQ ID NO:3, 5 and 7
CC (ADR29359, ADR29361 and ADR29363). Also described: (i) a peptide (II)
CC comprising the 3011 amino acid sequence of SEQ ID NO:1 (S1, ADR29357),
CC where in (S1) positions 2594-2602 are Phe, Ile, Leu, Trp, Pro, Glu, Asn,
CC Lys, and Val (SEQ ID NO:3), respectively or Phe, Leu, Pro, Trp, Gly, Ala,
CC Trp, Lys, and Val (SEQ ID NO:5) respectively; (2) a polynucleotide (III)
CC that encodes Phe-Leu-Ile-Trp-Pro-Glu-Asn-Lys-Val; Phe-Leu-Pro-Trp-Gly-Ala
CC -Trp-Lys-Val, or Ala-Leu-Tyr-Asp-Val-Thr-Lys-Leu (SEQ ID NO:3, 5 or 7);

CC (3) an antibody (IV) that recognises and binds to (I); (4) an immune
CC effector cell (V) that has been raised in vivo or in vitro or in the
CC presence and the expense of an antigen presenting cell that presents (I),
CC in the context of an MHC molecule; (5) a composition (VI) comprising (I),
CC of SEQ ID NO:3, 5 or 7 where (I) is individually characterised by an
CC ability to elicit an immune response against the same native ligand; (6)
CC a composition (VII) comprising (I) and a carrier; (7) a host cell (VIII)
CC comprising at least one or more (I) of SEQ ID NO:3, 5 or 7 that are
CC individually characterised by an ability to elicit an immune response
CC against the same native ligand; and (8) a composition comprising (VIII)
CC and a carrier. (I) have anti-HIV, virucide, antiinflammatory,
CC hepatotropic and immunostimulant activities, and can be used in vaccines.
CC (I) can be used for inducing an immune response in a subject which
CC involves delivering (I) to the subject. (I) is delivered in the context
CC of an MHC molecule. The MHC molecules presents (I) on the surface of an
CC antigen presenting cell. (I) is delivered as a polynucleotide that
CC encodes it. (IV) is useful for immunotherapy. (V) is useful for adoptive
CC immunotherapy. (VI) is useful for inducing an immune response in a
CC subject. (I) is useful as components of anti-HCV vaccines and to expand
CC immune effector cell that are specific for cells expressing HCV epitope.
CC (I) is useful for preparation of medicament for diagnosis and treatment
CC of diseases such as AIDS. The present sequence represents the hepatitis C
CC virus polypeptide precursor, which is used in the exemplification of the
CC present invention.

XX SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 8; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0;

Qy 18 GGVLAAALAAAYCLSL 30

Db 1664 GGVLAAALAAAYCLSL 1676

RESULT 272

ADX40785

ID ADX40785 standard; protein; 3011 AA.

AC ADX40785;

DT 21-APR-2005 (first entry)

DE HCV polymerase protein #8.

XX Immune stimulation; polymerase; enzyme.

OS Hepatitis C virus.

XX WO2005012502-A2.

XX 10-FEB-2005.

XX 29-MAR-2004; 2004WO-US009510.

XX 28-MAR-2003; 2003US-0458026P.

XX (EPIM-) EPIMUNE INC.

XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

XX WPI; 2005-132661/14.

XX Identifying a candidate peptide epitope, which induces a HLA class I CTL

PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.

XX Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the

CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX
SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 273
ADX40782
ID ADX40782 standard; protein; 3011 AA.

XX AC ADX40782;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #5.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.

XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (SPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.

XX PS Disclosure; Page 387-440; 458pp; English.
XX SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 274
ADX40784
ID ADX40784 standard; protein; 3011 AA.
XX AC ADX40784;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #7.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (SPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.

XX PS Disclosure; Page 388-440; 458pp; English.
XX SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 275
ADX40814
ID ADX40814 standard; protein; 3011 AA.

XX AC ADX40814;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #37.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.

XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (SPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.

XX PS Disclosure; Page 388-440; 458pp; English.
XX SQ Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 275
ADX40814
ID ADX40814 standard; protein; 3011 AA.

XX AC ADX40814;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #37.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.

PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Cheesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
SQ Sequence 3011 AA;
Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 1664 GGVLAALAAAYCLS 1676
RESULT 277
AEA06142
ID AEA06142 standard; protein; 3011 AA.
XX
AC AEA06142;
XX
DT 28-JUL-2005 (first entry)
XX
XX Full length consensus protein sequence of the hepatitis C virus genome.
XX
KW Hepatitis C virus infection; antiinflammatory; hepatotropic; virucide;
KW glycoprotein E1; glycoprotein E2; drug screening; antibody therapy;
KW liver cirrhosis; hepatocellular carcinoma; cytostatic; diagnostic;
KW diagnosis; gene therapy; nucleic acid vaccine.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 2999
FT /note= "Encoded by CTC"
XX
PN WO2005047481-A2.
XX
PD 26-MAY-2005.
XX
XX 09-NOV-2004; 2004WO-US037693.
PF
XX 12-NOV-2003; 2003US-0519536P.
PR
XX (PROG-) PROGENICS PHARM INC.
PA (YESH) UNIV YESHIV EINSTEIN COLLEGE.
XX
PI Cormier EG, Gardner J, Dragic T, Dumonceaux J;
XX
XX WPI; 2005-372370/38.
DR
XX N-ESDB; AEA06140.
XX
XX New modified nucleic acid comprising consecutive nucleotides having a
PT nucleotide sequence coding for a full length hepatitis C virus (HCV)
PT glycoprotein, useful in preparing a composition for treating or
PT preventing HCV infection.
XX
XX Example; SEQ ID NO 3; 152pp; English.
XX
XX The invention relates to sequences encoding hepatitis C virus (HCV)
CC glycoproteins. Envelope glycoproteins play a crucial role in viral
CC infectivity through their direct effect on processes including packaging
CC of virions, attachment to target cells, fusion and budding. In
CC particular, glycoproteins E1 and E2 are thought to be involved in
CC mediating virus entry into target cells. E1 and E2 are present as a non-
CC covalently associated heterodimer on the virus surface and undergo
CC modification by N-linked glycosylation. The new invention provides a
CC modified nucleic acid having a sequence encoding a full length hepatitis
CC C virus (HCV) glycoprotein consisting of E1 glycoprotein and E1/E2

PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Cheesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
SQ Sequence 3011 AA;
Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 1664 GGVLAALAAAYCLS 1676
RESULT 276
ADX40802
ID ADX40802 standard; protein; 3011 AA.
XX
AC ADX40802;
XX
DT 21-APR-2005 (first entry)
XX
XX HCV polymerase protein #25.
DE
XX Immune stimulation; polymerase; enzyme.
KW
XX Hepatitis C virus.
OS
XX WO2005012502-A2.
PN
XX 10-FEB-2005.
PD
XX 29-MAR-2004; 2004WO-US009510.
PF
XX 28-MAR-2003; 2003US-0458026P.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Baker DM, Livingston BD, Cheesnut RW, Sette A, Newman MJ;
PI WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide

CC glycoprotein heterodimer, which has at least one nucleotide alteration.
CC The alteration removes a RNA splice acceptor and RNA splice donor site
CC from the coding sequence. The elimination reduces an extent to which an
CC intron is excised from the coding sequence. The alteration comprises a
CC mutation in the HCV E1 coding sequence (SEQ ID NO: 2), so that a splice-
CC acceptor, or splice donor site is eliminated. Also disclosed is a method
CC of screening agents, to identify one that inhibits fusion; and an agent
CC that inhibits fusion or entry of HCV to a target cell capable of fusing
CC with the virus. The agent is an antibody, a peptide, a non-peptidyl agent
CC or a small molecule. Provided are methods of treating a subject with a
CC HCV-associated disorder; inhibiting the onset of a HCV-associated
CC hepatitis, liver cirrhosis and hepatocellular carcinoma. A diagnostic kit
CC is described comprising an antibody and instructions for detecting HCV in
CC human tissue. Treatment involves injecting a HCV glycoprotein or the
CC modified nucleic acid in a vector capable of expressing a HCV
CC glycoprotein. The present sequence is the full length amino acid
CC consensus sequence of the hepatitis C virus genome.

XX Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30

Db 1664 GGVLAALAAAYCLS 1676

RESULT 278

AE62080
ID AE62080 standard; protein; 3011 AA.

XX AC AE62080;

XX DT 25-AUG-2005 (first entry)

XX Hepatitis C virus strain H polyprotein.

XX Hepatitis C virus infection; antiinflammatory; hepatotropic; virucide;
XX gastrointestinal disease; drug screening.

XX Hepatitis C virus.

XX Key Location/Qualifiers

XX Region 1. .191

XX /label= Core

XX Region 192. .383

XX /label= E1

XX Region 384. .746

XX /label= E2

XX Region 747. .809

XX /label= P7

XX Region 810. .1026

XX /label= NS2

XX Region 1027. .1657

XX /label= NS3

XX Region 1658. .1711

XX /label= NS4A

XX Region 1712. .1972

XX /label= NS4B

XX Region 1973. .2420

XX /label= NS5A

XX Region 2421. .3011

XX /label= NS5B

XX WO2005053516-A2.

XX 16-JUN-2005.

XX 01-DEC-2004; 2004WO-US040120.

PR 01-DEC-2003; 2003US-0525989P.

XX (TEXA) UNIV TEXAS.

XX Lemon SM, Yi M;

XX WPI; 2005-418069/42.

DR N-PSDB; AEA62079.

XX New replication competent polynucleotide comprising a 5' non-translated
XX region (NTR), a 3' NTR, and a first coding sequence present between the
XX 5' NTR and 3' NTR and encoding a hepatitis C virus polyprotein, useful in
XX drug discovery.

XX Disclosure; SEQ ID NO 4; 102pp; English.

XX The present sequence is that of the hepatitis C virus (HCV) strain H
XX polyprotein. The invention provides replication competent polynucleotides
XX that include a 5' non-translated region (NTR), a 3' NTR and a coding
XX sequence encoding an HCV polyprotein that comprises an Ile residue at
XX about amino acid 2204, and an adaptive mutation selected from Arg at
XX about amino acid 1067, Arg at about amino acid 1691, Val at about amino
XX acid 2080, Ile at about amino acid 1655, Arg at about amino acid 2040,
XX and Arg at about amino acid 1188, or a combination of these. The
XX polyprotein may include the cleavage products core, E1, E2, P7, NS2, NS3,
XX NS4A, NS4B, NS5A and NS5B. The polynucleotide may further comprise a
XX second coding sequence encoding a marker or transactivator, or a
XX nucleotide sequence having cis-acting ribozyme activity located 3' of the
XX 3' NTR. Also provided are: a method for making the replication component
XX polynucleotide; a method for using the replication competent
XX polynucleotide to identify a compound that inhibits replication of the
XX replication competent nucleotide; a method for selecting a replication
XX competent polynucleotide; and a method for detecting a replication
XX competent polynucleotide.

XX Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30

Db 1664 GGVLAALAAAYCLS 1676

RESULT 279

AE62088

ID AE62088 standard; protein; 3011 AA.

XX AC AE62088;

XX DT 25-AUG-2005 (first entry)

XX Hepatitis C virus genotype 1a polyprotein.

XX Hepatitis C virus infection; antiinflammatory; hepatotropic; virucide;
XX gastrointestinal disease; drug screening.

XX Hepatitis C virus.

XX Key Location/Qualifiers

XX Region 1. .191

XX /label= Core

XX Region 192. .383

XX /label= E1

XX Region 384. .746

XX /label= E2

XX Region 747. .809

XX /label= P7

XX Region 810. .1026

XX /label= NS2

XX Region 1027. .1657

```

FT      Region      /label= NS3
FT      1658..1711
FT      /label= NS4A
FT      1712..1972
FT      /label= NS4B
FT      1973..2420
FT      /label= NS5A
FT      2421..3011
FT      /label= NS5B
XX
XX      WO2005053516-A2.
XX
XX      16-JUN-2005.
XX
XX      01-DEC-2004; 2004WO-US040120.
XX
XX      01-DEC-2003; 2003US-0525989P.
XX
XX      (TEXA ) UNIV TEXAS.
XX
XX      Lemon SM, Yi M;
XX
XX      WPI; 2005-418069/42.
XX      N-PSDB; AEA62087, AEA62091.
XX
XX      New replication competent polynucleotide comprising a 5' non-translated
XX      region (NTR), a 3' NTR, and a first coding sequence present between the
XX      5' NTR and 3' NTR and encoding a hepatitis C virus polyprotein, useful in
XX      drug discovery.
XX
XX      Disclosure; SEQ ID NO 12; 102pp; English.
XX
XX      The present sequence is that of the hepatitis C virus (HCV) genotype 1a
XX      polyprotein. The invention provides replication competent polynucleotides
XX      that include a 5' non-translated region (NTR), a 3' NTR and a coding
XX      sequence encoding an HCV polyprotein that comprises an Ile residue at
XX      about amino acid 2204, and an adaptive mutation selected from Arg at
XX      about amino acid 1067, Arg at about amino acid 1691, Val at about amino
XX      acid 2080, Ile at about amino acid 1655, Arg at about amino acid 2040,
XX      and Arg at about amino acid 1188, or a combination of these. The
XX      polyprotein may include the cleavage products core, E1, E2, P7, NS2, NS3,
XX      NS4A, NS4B, NS5A and NS5B. The polynucleotide may further comprise a
XX      second coding sequence encoding a marker or transactivator, or a
XX      nucleotide sequence having cis-acting ribozyme activity located 3' of the
XX      3' NTR. Also provided are: a method for making the replication component
XX      polynucleotide; a method for using the replication component
XX      polynucleotide to identify a compound that inhibits replication of the
XX      replication competent nucleotide; a method for selecting a replication
XX      competent polynucleotide; and a method for detecting a replication
XX      competent polynucleotide.
XX      Sequence 3011 AA;

  Query Match      11.0%; Score 13; DB 9; Length 3011;
  Best Local Similarity 100.0%; Pred. No. 0.006;
  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      18 GGVLAAALAAAYCIS 30
      |||||
Db      1664 GGVLAAALAAAYCIS 1676

RESULT 280
AEA62078
ID      AEA62078 standard; protein; 3011 AA.
XX
XX      AEA62078;
XX
XX      25-AUG-2005 (first entry)
XX
XX      Hepatitis C virus strain H77 polyprotein.
XX
XX      Hepatitis C virus infection; antiinflammatory; hepatotropic; virucide;

```

gastrointestinal disease; drug screening.

Hepatitis C virus.

Key Location/Qualifiers

Region 1..191

Region /label= Core

Region 192..383

Region /label= E1

Region 384..746

Region /label= E2

Region 747..809

Region /label= P7

Region 810..1026

Region /label= NS2

Region 1027..1657

Region /label= NS3

Region 1658..1711

Region /label= NS4A

Region 1712..1972

Region /label= NS4B

Region 1973..2420

Region /label= NS5A

Region 2421..3011

Region /label= NS5B

WO2005053516-A2.

16-JUN-2005.

01-DEC-2004; 2004WO-US040120.

01-DEC-2003; 2003US-0525989P.

(TEXA) UNIV TEXAS.

Lemon SM, Yi M;

WPI; 2005-418069/42.

N-PSDB; AEA62077.

New replication competent polynucleotide comprising a 5' non-translated

region (NTR), a 3' NTR, and a first coding sequence present between the

5' NTR and 3' NTR and encoding a hepatitis C virus polyprotein, useful in

drug discovery.

Disclosure; SEQ ID NO 2; 102pp; English.

The present sequence is that of the hepatitis C virus (HCV) strain H77

polyprotein. The invention provides replication competent polynucleotides

that include a 5' non-translated region (NTR), a 3' NTR and a coding

sequence encoding an HCV polyprotein that comprises an Ile residue at

about amino acid 2204, and an adaptive mutation selected from Arg at

about amino acid 1067, Arg at about amino acid 1691, Val at about amino

acid 2080, Ile at about amino acid 1655, Arg at about amino acid 2040,

and Arg at about amino acid 1188, or a combination of these. The

polyprotein may include the cleavage products core, E1, E2, P7, NS2,

NS4A, NS4B, NS5A and NS5B. The polynucleotide may further comprise a

second coding sequence encoding a marker or transactivator, or a

nucleotide sequence having cis-acting ribozyme activity located 3' of the

3' NTR. Also provided are: a method for making the replication component

polynucleotide; a method for using the replication component

polynucleotide to identify a compound that inhibits replication of the

replication competent nucleotide; a method for selecting a replication

competent polynucleotide; and a method for detecting a replication

competent polynucleotide.

Sequence 3011 AA;

Query Match 11.0%; Score 13; DB 9; Length 3011;

Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 281
AEB17053
ID AEB17053 standard; protein; 3011 AA.
XX AEB17053;
XX
DT 22-SEP-2005 (first entry)
XX
DE Hepatitis C virus MKO-Z protein, SEQ ID NO: 20.
XX
KW RNA virus; microorganism; hepatitis C virus infection; antinflammatory;
KW hepatotropic; virucide; gastrointestinal disease; infection.
XX
OS Hepatitis C virus.
XX

Key Location/Qualifiers
Region 1..191
FT /note= "Core protein"
Region 192..383
FT /note= "Envelope protein (E1)"
Region 384..746
FT /note= "Envelope protein (E2)"
Region 747..809
FT /note= "p7 protein"
Region 810..1027
FT /note= "NS2 protein"
Region 1028..1657
FT /note= "NS3 protein"
Region 1658..1711
FT /note= "NS4A protein"
Region 1712..1982
FT /note= "NS4B protein"
Region 1983..2420
FT /note= "NS5A protein"
Region 2421..3011
FT /note= "NS5B protein"

US2005153281-A1.
14-JUL-2005.
06-DEC-2004; 2004US-00006313.
23-DEC-1999; 99US-0171909P.
23-DEC-2000; 2000US-00747419.
27-SEP-2001; 2001US-0325236P.
13-NOV-2001; 2001US-0338123P.
27-SEP-2002; 2002US-00259275.
(TEXA) UNIV TEXAS SYSTEM.
Lemon SM, Yi M;
WPI; 2005-496820/50.
N-PSDB; AEB17050.
Detecting replication competent hepatitis C virus (HCV) RNA by incubating a vertebrate cell comprising an HCV RNA and detecting the detectable marker, where its presence indicates the cell comprises a replication competent HCV RNA.
Disclosure; SEQ ID NO 20; 142pp; English.
The present invention provides a replication competent hepatitis C virus (HCV) RNA that includes a heterologous polynucleotide. The invention also includes methods for modifying a HCV polynucleotide, selecting a replication competent HCV polynucleotide, detecting a replication competent HCV polynucleotide and identifying a compound that inhibits

CC replication of a HCV polynucleotide. The present sequence is hepatitis C
CC virus (HCV) MKO-Z (full-length modified cDNA) protein. This sequence
CC includes 5' non-translated RNA (5' NTR), core protein, envelope
CC glycoproteins (E1 and E2) and six non-structural replicative proteins
CC which includes NS2 (which demonstrates cis active metalloproteinase
CC activity at the NS2/NS3 cleavage site), NS3 (a serine
CC proteinase/NTase/RNA helicase), NS4A (serine proteinase accessory
CC factor), NS4B, NS5A and NS5B (RNA-dependent RNA polymerase). This
CC sequence is encoded by a portion (corresponding to nucleotides 342 to
CC 10803) of hepatitis C virus MKO-Z (full length modified cDNA) protein.
XX
SQ Sequence 3011 AA;
Query Match 11.0%; Score 13; DB 9; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 282
AAU99289
ID AAU99289 standard; protein; 3012 AA.
XX
AC AAU99289;
XX
DT 24-SEP-2002 (first entry)
XX
DE Hepatitis C virus (HCV) H77 consensus protein.
XX
KW Hepatitis C virus; HCV; productive replication; infection; antiviral;
KW H77; consensus protein.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 3012
FT /note= "Encoded by TGA"
XX
PN US6392028-B1.
XX
PD 21-MAY-2002.
XX
PP 04-MAR-1998; 98US-00034756.
XX
PR 04-MAR-1997; 97US-0039843P.
XX 04-MAR-1997; 97US-00811566.
XX (UNIW) UNIV WASHINGTON.
XX
PI Rice CM, Kolykhalov AA;
DR WPI; 2002-478540/51.
DR N-PSDB; ABK87285.
XX
PT New DNA or RNA comprising a hepatitis C virus sequence that contains 3'
PT and 5' non-translated regions flanking a polyprotein coding region,
PT useful for detecting and developing treatment for infection by hepatitis
PT C virus.
XX
PS Disclosure; Col 75-90; 114pp; English.
XX
CC The present invention relates to a new DNA or RNA comprising a hepatitis
CC C virus (HCV) sequence capable of productive replication in a host cell.
CC The molecules of the invention comprise 5' to 3' on the positive sense 5'
CC strand a functional HCV 5' non-translated region comprising an extreme 5'
CC -terminal conserved sequence, an HCV polyprotein coding region and a
CC functional HCV 3' non-translated region comprising an extreme 3'-terminal
CC conserved region. The DNA or RNA is used to detect and develop treatment
CC for infection by hepatitis C virus. The present amino acid sequence
CC represents the hepatitis C virus (HCV) H77 consensus protein

XX	Sequence 3012 AA;
SQ	
	Query Match 11.0%; Score 13; DB 5; Length 3012; Best Local Similarity 100.0%; Pred.No. 0.006; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	18 GGVLAAALAAAYCLS 30
Dd	1664 GGVLAAALAAAYCLS 1676
	RESULT 283
ABU61848	
ID	ABU61848 standard; protein; 3012 AA.
XX	
AC	ABU61848;
XX	
DT	15-AUG-2003 (first entry)
XX	
DE	HCV H77 consensus sequence.
XX	
KW	HCV H77; gene therapy; vaccine; HCV infection; hepatitis C virus.
XX	
OS	Hepatitis C virus.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 3012
PT	/note= "Encoded by TGA"
XX	
PN	US2003028010-A1.
XX	
PD	06-FEB-2003.
XX	
PP	28-NOV-2001; 2001US-00995937.
XX	
PR	04-MAR-1997; 97US-0039843P.
PR	04-MAR-1998; 98US-00034756.
XX	
PA	(UNIW) UNIV WASHINGTON.
XX	
PI	Rice CM, Kolykhalov AA;
XX	
DR	WFI; 2003-466160/44.
N-PSDB; ACA62466.	
XX	
PT	Novel genetically engineered hepatitis C virus, for identifying animal permissive for HCV infection, has positive sense nucleic acid, 5'- and 3'-non-translated regions, open reading frame encoding HCV polyprotein.
PS	Disclosure; Page 41-48; 120pp; English.
CC	The invention relates to a genetically engineered hepatitis C virus (HCV) nucleic acid clone. The genetically engineered HCV is useful for identifying a cell line or animal that is permissive for infection with HCV. The genetically engineered HCV is useful selecting for HCV with adaptive mutations that permit higher levels of HCV replication in a permissive cell line. The genetically engineered HCV DNA or RNA is useful for infecting an animal with HCV and for propagating HCV in vitro. The genetically engineered HCV or HCV DNA or RNA is useful for transducing an animal susceptible to HCV infection with a heterologous gene. A host cell line, a non-human animal or an in vitro cell line infected with the HCV genomic RNA of the invention is useful for producing HCV virus particles and for screening for agents capable of modulating HCV replication. The genetically engineered HCV is useful for producing HCV particle proteins. A HCV virus particle comprising a replication-competent or replication-defective HCV genome RNA is useful for producing antibodies to HCV. A HCV virus particle comprising a replication-competent or replication-defective HCV genome RNA is useful in an in vitro method for detecting antibodies to HCV in a biological sample from a subject. The genetically engineered HCV is useful for therapeutically, vaccine and diagnostic applications. A plasmid encoding the genetically engineered HCV is useful for gene therapy and in gene

Query Match 11.0%; Score 13; DB 4; Length 3015;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1668 GGVLAALAAAYCLS 1680

RESULT 285
AAB30731
ID AAB30731 standard; protein; 3015 AA.
AC AAB30731;
DT 02-APR-2001 (first entry)
DE Amino acid sequence of chimeric Hepatitis C virus clone pH77(p7)CV-J6S.
KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
OS Synthetic.
OS Hepatitis C virus.
FN WO200075338-A2.
PD 14-DEC-2000.
PF 02-JUN-2000; 2000WO-US015446.
PR 04-JUN-1999; 99US-0137693P.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
DR WPI; 2001-061728/07.
DR N-PSDB; AAC86646.
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for developing vaccines, for diagnosis of hepatitis C virus and in screening assays for identification of antiviral agents.
XX Disclosure; Page 119-130; 167pp; English.
XX AAB30730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC-J6CH or HCV strain pCV-H77C, and the NS genes were derived from HCV strain pCV-H77C. Such HCV sequences are capable of expressing the virus when transfected into cells. The HCV protein is useful for assaying candidate antiviral agents for activity against HCV. Antibodies specific for HCV polypeptide are useful in prevention and treatment of diseases caused by HCV in animals, in particular humans. The HCV polypeptides serve as immunogens in the development of vaccines for preventing HCV in mammals or as antigens in diagnostic assays for detecting the presence of HCV in biological samples. The HCV polynucleotide is also useful for identifying cell lines capable of supporting the replication of HCV in vitro and to produce attenuated viral strains via passage in vitro or in vivo

Query Match 11.0%; Score 13; DB 4; Length 3015;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1668 GGVLAALAAAYCLS 1680

RESULT 286
AAB30730
ID AAB30730 standard; protein; 3015 AA.
AC AAB30730;
DT 02-APR-2001 (first entry)
DE Amino acid sequence of chimeric Hepatitis C virus clone pH77CV-J6S.
KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
OS Synthetic.
OS Hepatitis C virus.
FN WO200075338-A2.
PD 14-DEC-2000.
PF 02-JUN-2000; 2000WO-US015446.
PR 04-JUN-1999; 99US-0137693P.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Yanagi M, Bukh J, Emerson SU, Purcell RH;
DR WPI; 2001-061728/07.
DR N-PSDB; AAC86645.
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for developing vaccines, for diagnosis of hepatitis C virus and in screening assays for identification of antiviral agents.
XX Disclosure; Page 103-115; 167pp; English.
XX AAB30730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC-J6CH or HCV strain pCV-H77C, and the NS genes were derived from HCV strain pCV-H77C. Such HCV sequences are capable of expressing the virus when transfected into cells. The HCV protein is useful for assaying candidate antiviral agents for activity against HCV. Antibodies specific for HCV polypeptide are useful in prevention and treatment of diseases caused by HCV in animals, in particular humans. The HCV polypeptides serve as immunogens in the development of vaccines for preventing HCV in mammals or as antigens in diagnostic assays for detecting the presence of HCV in biological samples. The HCV polynucleotide is also useful for identifying cell lines capable of supporting the replication of HCV in vitro and to produce attenuated viral strains via passage in vitro or in vivo

Query Match 11.0%; Score 13; DB 4; Length 3015;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1668 GGVLAALAAAYCLS 1680

RESULT 287
AAB30732
ID AAB30732 standard; protein; 3015 AA.
AC AAB30732;
DT 02-APR-2001 (first entry)
DE Amino acid sequence of chimeric Hepatitis C virus clone J6S.
XX

KW HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
XX Synthetic.
OS Hepatitis C virus.
XX WO200075338-A2.
PN 14-DEC-2000.
PD 02-JUN-2000; 2000WO-US015446.
PF 04-JUN-1999; 99US-0137693P.
XX (USHS) US DEPT HEALTH & HUMAN SERVICES.
PA Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX WPI; 2001-061728/07.
DR N-PSDB; AAC86647.
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
PT assays for identification of antiviral agents.
XX Disclosure; Page 134-146; 167pp; English.
XX AAB30730-33 are encoded by chimeric cDNA clones of infectious Hepatitis C
CC virus (HCV). In each clone the C, E1 and E2 genes are derived from HCV
CC strain HC-J6CH genotype 2a. The p7 protein was derived from HCV strain HC
CC -J6CH or HCV strain PCV-H77C, and the NS genes were derived from HCV
CC strain PCV-H77C. Such HCV sequences are capable of expressing the virus
CC when transfected into cells. The HCV protein is useful for assaying
CC candidate antiviral agents for activity against HCV. Antibodies specific
CC for HCV polypeptide are useful in prevention and treatment of diseases
CC caused by HCV in animals, in particular humans. The HCV polypeptides
CC serve as immunogens in the development of vaccines for preventing HCV in
CC mammals or as antigens in diagnostic assays for detecting the presence of
CC HCV in biological samples. The HCV polynucleotide is also useful for
CC identifying cell lines capable of supporting the replication of HCV in
CC vitro and to produce attenuated viral strains via passage in vitro or in
XX vivo
XX Sequence 3015 AA;
Qy Query Match 11.0%; Score 13; DB 4; Length 3015;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 1668 GGVLAALAAAYCLS 1680
|||||
RESULT 288
ABG73195
ID ABG73195 standard; protein; 3180 AA.
XX AC ABG73195;
XX DT 11-APR-2003 (first entry)
XX DE MKO-Z viral polypeptides #1.
XX KW Replication competent; hepatitis C virus; HCV; 3' non-translated RNA;
KW 3'NTR; chronic viral hepatitis; hepatic fibrosis; cirrhosis;
KW hepatocellular carcinoma; MKO-Z.
XX OS Hepatitis c virus.
XX FH Key Location/Qualifiers
FT Protein 1..191
FT Product= "Core protein"
FT Protein 192..383

FT Protein
FT 384..746
FT Product= "Envelope protein 2"
FT 747..809
FT Product= "Envelope protein 2-p7"
FT 810..1027
FT Product= "Non-structural protein 2"
FT 1028..1657
FT Product= "Non-structural protein 3"
FT 1658..1711
FT Product= "Non-structural protein 4A"
FT 1712..1972
FT Product= "Non-structural protein 4B"
FT 1973..2420
FT Product= "Non-structural protein 5A"
FT 2421..3011
FT Product= "Non-structural protein 5B"
FT Misc-difference 3012
FT /note= "In-frame stop codon encoded by TGA"
FT Misc-difference 3016
FT /note= "In-frame stop codon encoded by TAA"
FT Misc-difference 3049
FT /note= "In-frame stop codon encoded by TAG"
FT Misc-difference 3152
FT /note= "In-frame stop codon encoded by TGA"
XX US2002155582-A1.
PN 24-OCT-2002.
XX 23-DEC-2000; 2000US-00747419.
XX 23-DEC-1999; 99US-0171909P.
XX (LEMO/) LEMON S M.
XX (YIMW/) YI M.
XX Lemon SM, Yi M;
XX WPI; 2003-182640/18.
XX N-PSDB; ABX10617.
XX Novel replication competent hepatitis C virus for producing infectious
PT viral particles and as antigen for detecting hepatitis C virus
PT antibodies, comprises hepatitis C virus genome and heterologous
PT polynucleotide.
XX Disclosure; Fig 10; 37pp; English.
XX The invention discloses a replication competent hepatitis C virus (HCV)
CC comprising a HCV virus genome and a heterologous polynucleotide, where
CC the HCV genome comprises a 3' non-translated RNA and the heterologous
CC polynucleotide is present in the 3' non-translated RNA. HCV is a cause of
CC chronic viral hepatitis, hepatic fibrosis, cirrhosis and/or the
CC development of hepatocellular carcinoma. A cell comprising the HCV is
CC useful for selecting or detecting a replication competent HCV, for
CC identifying a compound that inhibits replication of HCV, for producing
CC infectious viral particles which are useful as a source of virus
CC particles for various assays, including evaluating methods for
CC inactivating particles, excluding particles from serum, identifying a
CC neutralising compound and as an antigen for use in detecting anti-HCV
CC antibodies in an animal. The cell comprising the HCV is also useful for
CC identifying a variant HCV. An HCV particle is useful as an antigen, as a
CC positive-control in assays that test for the presence of anti-HCV
CC antibodies, to produce antibodies to detect the presence of viral
CC particles in biological samples (e.g. blood products and cell-free blood
CC products) and as a source of viral antigen to measure the presence and
CC amount of antibody present in an animal. The sequence presented is the
CC viral polypeptides encoded by the modified HCV (MKO-Z) polynucleotide
XX Sequence 3180 AA;
SQ Query Match 11.0%; Score 13; DB 6; Length 3180;

Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 1664 GGVLAALAAAYCLS 1676

RESULT 289
ADD67948
ID ADD67948 standard; protein; 3208 AA.
AC ADD67948;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus polyprotein seq id 20.
XX
KW antiviral; hepatitis C virus; HCV; viral replication inhibitor;
KW replication competent HCV; 3' non-translated region; polyprotein.
XX
OS Hepatitis C virus.
XX
PN US2003125541-A1.
XX
PD 03-JUL-2003.
XX
XX 27-SEP-2002; 2002US-00259275.
XX
XX 23-DEC-1999; 99US-0171909P.
XX 23-DEC-2000; 2000US-00747419.
XX 21-SEP-2001; 2001US-0325236P.
XX 13-NOV-2001; 2001US-0338123P.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Lemon SM, Yi M;
XX
XX WPI; 2003-811006/76.
XX
XX Identifying a compound that inhibits replication of a hepatitis C virus
XX (HCV) RNA comprises contacting a cell comprising a replication competent
XX HCV RNA containing a heterologous polynucleotide encoding a
XX transactivator, with a compound.
XX
XX Disclosure; SEQ ID NO 20; 95pp; English.
XX
XX The invention describes a method of identifying a compound that inhibits
XX replication of a hepatitis C virus (HCV) RNA. The method comprises
XX contacting a cell comprising a replication competent HCV RNA containing a
XX heterologous polynucleotide having a first coding sequence encoding a
XX transactivator, with a compound. The method is useful for identifying a
XX compound that inhibits replication of HCV RNA. The kit is useful for
XX detecting replication competent HCV RNA. This is the amino acid sequence
XX of hepatitis C virus polyprotein.
XX
SQ Sequence 3208 AA;

Query Match 11.0%; Score 13; DB 7; Length 3208;
Best Local Similarity 100.0%; Pred. No. 0.0064;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 1664 GGVLAALAAAYCLS 1676

RESULT 290
ADV23776
ID ADV23776 standard; protein; 3261 AA.
XX
AC ADV23776;
XX

DT 10-MAR-2005 (first entry)
XX
DE Hepatitis C virus strain H77 protein.
XX
KW vaccine; virucide; antigen; autoimmune disease; infection;
KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
KW hodgkin's lymphoma.
XX
OS Hepatitis C virus strain H77.
XX
PN WO2004108753-A1.
XX
PD 16-DEC-2004.
XX
XX 10-JUN-2004; 2004WO-AU000775.
XX
XX 10-JUN-2003; 2003AU-00902875.
XX 25-MAR-2004; 2004AU-00901589.
XX
XX (UYME) UNIV MELBOURNE.
XX
XX Kent SJ;
XX
XX WPI; 2005-031657/03.
XX
XX Use of at least one set of peptides in the preparation of a medicament
XX for modulating an immune response, and for treating cancer or yeast,
XX viral, bacterial, protozoal and mycoplasma infections.
XX
XX Disclosure; SEQ ID NO 2196; 645pp; English.
XX
XX The invention relates to the use of at least one set of peptides in the
XX preparation of a medicament for modulating an immune response, where
XX individual peptides of a respective set comprise different portions of an
XX amino acid sequence corresponding to a single polypeptide of interest and
XX display partial sequence identity or similarity to at least one other
XX peptide of the same set of peptides (i.e. they are overlapping). Also
XX included are an antigen-presenting cell which has been contacted with the
XX peptides above and thus presents the peptides, a population of such
XX antigen-presenting cells, a process for producing antigen-presenting
XX cells for modulating an immune response to a polypeptide of interest, a
XX method for producing antigen-specific lymphocytes, a composition
XX comprising at least one set of the peptides (and a carrier and/or
XX diluent), a method for modulating an immune response to a polypeptide of
XX interest comprising administering to a patient in need at least one set
XX of the peptides, a method for treatment and/or prophylaxis of a disease
XX or condition associated with the presence of a polypeptide of interest
XX and a composition of matter for modulating an immune response in a
XX subject to a target antigen. The polypeptide of interest is also a
XX disease- or condition-associated polypeptide that is a polypeptide
XX produced by a pathogenic organism or a cancer, and produced by a
XX pathogenic organism selected from yeast, viruses, bacteria, helminths,
XX protozoans and mycoplasmas. The disease- or condition-associated
XX polypeptide is produced by a cancer selected from melanoma, lung cancer,
XX breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
XX cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
XX lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
XX antigen-presenting cells or their precursors are useful in the
XX preparation of a medicament for the treatment of a disease or condition
XX in a subject, which disease or condition is associated with the presence
XX of aberrant expression of a target antigen, where the antigen-presenting
XX cells or their precursors have not been subjected to activating
XX conditions but have been contacted with an antigen that corresponds to
XX the target antigen to express a processed or modified form of the antigen
XX for presentation to the subject's immune system. The present sequence is
XX a viral protein from which a set of overlapping immunogenic peptides was
XX derived.
XX
SQ Sequence 3261 AA;

Query Match 11.0%; Score 13; DB 9; Length 3261;

Best Local Similarity 100.0%; Pred. No. 0.0065; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
|||||

Db 1812 GGVLAALAAAYCLIS 1824

RESULT 291

AAU84799
ID AAU84799 standard; protein; 5985 AA.

XX AC AAU84799;

XX DT 08-MAY-2002 (first entry)

XX DE HCV HepC1a scrambled.

XX KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX KW viral infection; human immunodeficiency virus; melanoma;
XX KW bacterial infection; Salmonella; Legionella; parasitic infection;
XX KW Trypanosoma; Toxoplasma; Giardia.

XX OS Hepatitis C virus.

XX PN WO200190197-A1.

XX PD 29-NOV-2001.

XX PF 25-MAY-2001; 2001WO-AU000622.

XX PR 26-MAY-2000; 2000AU-00007761.

XX PA (AUSU) UNIV AUSTRALIAN NAT.

XX PI Thomson SA, Ramshaw IA;

XX DR WPI; 2002-147575/19.

XX DR N-PSDB; ABK36637.

XX PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.

XX PS Example 2; Fig 26; 364pp; English.

XX CC The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a cassette protein consisting of several peptides derived from a parent
CC protein. One or more cassettes are used to construct a savine of the
CC invention

XX SQ Sequence 5985 AA;

Query Match 11.0%; Score 13; DB 5; Length 5985;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
|||||

Db 1591 GGVLAALAAAYCLIS 1603

RESULT 292

AAJ03066
ID AAJ03066 standard; peptide; 15 AA.

XX AC AAJ03066;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #3057.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccines against hepatitis C virus.

XX PS Disclosure; Page 174; 214pp; English.

XX CC The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention

XX SQ Sequence 15 AA;

Query Match 10.2%; Score 12; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||

Db 1 GGVLAALAAAYCL 12

RESULT 293

AAJ03372
ID AAJ03372 standard; peptide; 15 AA.

XX AC AAJ03372;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #3363.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.

```
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX XX
XX PD 29-MAR-2001.
XX XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX XX
XX PR 19-JUL-1999; 99US-00357737.
XX XX
XX PA (EPIM-) EPIMMUNE INC.
XX XX
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX XX
XX DR WPI; 2001-308046/32.
XX XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 177; 214pp; English.
XX XX
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX XX
XX SQ Sequence 15 AA;

Query Match 10.2%; Score 12; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCL 29
Db 2 GGVLAALAAAYCL 13

RESULT 295
AAJ03246
ID AAJ03246 standard; peptide; 15 AA.
XX
XX AC AAJ03246;
XX XX
XX DT 02-JUL-2001 (first entry)
XX XX
XX DE Hepatitis C virus epitope #3237.
XX XX
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX XX
XX OS Hepatitis C virus.
XX XX
XX PN WO200121189-A1.
XX XX
XX PD 29-MAR-2001.
XX XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX XX
XX PR 19-JUL-1999; 99US-00357737.
XX XX
XX PA (EPIM-) EPIMMUNE INC.
XX XX
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX XX
XX DR WPI; 2001-308046/32.
XX XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 176; 214pp; English.
XX XX
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX XX
XX SQ Sequence 15 AA;

Query Match 10.2%; Score 12; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00063;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCL 29
Db 1 GGVLAALAAAYCL 12

RESULT 294
AAJ03508
ID AAJ03508 standard; peptide; 15 AA.
XX
XX AC AAJ03508;
XX XX
XX DT 02-JUL-2001 (first entry)
XX XX
XX DE Hepatitis C virus epitope #3499.
XX XX
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX XX
XX OS Hepatitis C virus.
XX XX
XX PN WO200121189-A1.
XX XX
XX PD 29-MAR-2001.
XX XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX XX
XX PR 19-JUL-1999; 99US-00357737.
XX XX
XX PA (EPIM-) EPIMMUNE INC.
XX XX
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX XX
XX DR WPI; 2001-308046/32.
XX XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 178; 214pp; English.
XX XX
```

XX 06-AUG-2003 (revised)
 DT 22-JUN-2001 (first entry)
 XX HCV antigen, NS4 1655-1675.
 DE Antigen; immunostimulant; vaccine; pharmaceutical composition; antiviral;
 XX viral infection.
 KW Hepatitis C virus.
 OS WO200124822-A2.
 XX 12-APR-2001.
 PD 02-OCT-2000; 2000WO-EP009657.
 XX 01-OCT-1999; 99AT-00001680.
 PR (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
 XX Fleitmann J, Mattner F, Buschle M, Melling J;
 FI WPI; 2001-290577/30.
 DR New pharmaceutical composition comprising an antigen, an
 XX immunostimulating substance and a polycationic polymer, useful in
 PT manufacturing vaccines.
 XX Claim 12; Page 16; 20pp; English.
 PS The present invention relates to a pharmaceutical composition comprising
 CC (a) an antigen; (b) an immunostimulating substance consisting of
 CC neuroactive compounds, hormones, compounds having growth hormone activity
 CC or their mixtures; and (c) a polycationic polymer. The present sequence
 CC is an antigenic peptide derived from Hepatitis C virus, which was used in
 CC the present invention. The composition is useful in manufacturing
 CC vaccines. (Updated on 06-AUG-2003 to correct OS field.)
 XX Sequence 21 AA;
 SQ

Query Match 10.2%; Score 12; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.00084;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 18 GGVLAALAAAYCL 29
 DB 10 GGVLAALAAAYCL 21

RESULT 297
 AA21988
 ID AA21988 standard; peptide; 21 AA.
 XX AA21988;
 AC 25-JUL-2002 (first entry)
 DT Hepatitis C virus epitope #2 specific for CD4+ T-lymphocytes.
 XX Hepatitis C virus; HCV; epitope; CD4+ T-lymphocyte; infection; vaccine;
 DE virucide; antiinflammatory; hepatotropic.
 KW Hepatitis C virus.
 OS WO200226785-A2.
 XX 04-APR-2002.
 PD 28-SEP-2001; 2001WO-EP011263.
 XX 28-SEP-2000; 2000EP-00121138.
 XX

PA (IMMU-) IMMUSYSTEMS GMBH.
 XX Gerlach JT, Diepolder H;
 PI WPI; 2002-362486/39.
 DR New hepatitis C virus epitopes which are specific for CD40-positive T
 XX lymphocytes, useful for diagnosis and vaccination of hepatitis C virus
 PT infection.
 PT Claim 1; Page 14; 34pp; German.
 PS The present invention relates to Hepatitis C virus (HCV) epitopes that
 XX are specific for CD4+ T-lymphocytes and their derivatives with comparable
 CC specificity. HCV epitopes of the invention are useful for diagnosis of
 CC HCV infection (by determining the level of CD4+ cells specific for a
 CC particular epitope) and particularly when formulated as a vaccine, for
 CC treatment and prevention of infection. DNAs that encode HCV epitopes can
 CC also be used in vaccines. The present sequence is a HCV epitope specific
 CC for CD4+ T-lymphocytes
 XX Sequence 21 AA;
 SQ

Query Match 10.2%; Score 12; DB 5; Length 21;
 Best Local Similarity 100.0%; Pred. No. 0.00084;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 18 GGVLAALAAAYCL 29
 DB 10 GGVLAALAAAYCL 21

RESULT 298
 AAR82855
 ID AAR82855 standard; protein; 54 AA.
 XX AAR82855;
 AC 25-MAR-2003 (revised)
 DT 09-APR-1996 (first entry)
 XX NS4A protein.
 DE NS3; serine protease; hepatitis C virus; HCV; NS4A; therapy.
 KW Hepatitis C virus.
 XX OS
 XX WO9522985-A1.
 PN 31-AUG-1995.
 PD 14-FEB-1995; 95WO-IT000018.
 PF 23-FEB-1994; 94IT-RM000092.
 XX (RICE-) IST RICERCHÉ BIOL MOLECOLARE ANGELETTI.
 PA De Francesco R, Failla C, Tomei L;
 XX WPI; 1995-311381/40.
 DR In vitro reproduction of hepatitis C virus NS3 protease activity - by
 PT including the NS4A cofactor in the mixt., useful for screening cpds. that
 PT inhibit NS3.
 XX Claim 6; Page 18-19; 26pp; English.
 PS This sequence represents the Hepatitis C virus (HCV) NS4A protein. This
 CC sequence is a cofactor of the NS3 serine protease domain (see AAR82854).
 CC Optimal serine protease activity is obtained when NS4A and NS3 are
 CC present in a 1:1 ratio. The cleavage site between these two proteins on
 CC the HCV genome can be mutated so that the components remain covalently
 CC bonded. These sequences are included in a composition that can be used in

CC an assay system. This assay system can be used to select compounds
CC that inhibit NS3 activity, e.g. potential therapeutic agents. NS4A can be
CC used for screening enzyme inhibitors. (Updated on 25-MAR-2003 to correct
CC PR field.)
XX
XX

SQ Sequence 54 AA;

Query Match 10.2%; Score 12; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 7 GGVLAALAAAYCL 18

RESULT 299

AAW37808
ID AAW37808 standard; peptide; 54 AA.

AC AAW37808;

XX 19-AUG-1998 (first entry)

XX Nonstructural domain protein 4A of Hepatitis C virus.

XX Nonstructural protein 3; NS3; HCV; detergent-free NS3 protease;
KW screening assay; inhibition; protease activity; diagnosis; HCV infection;
KW NS4A.
XX

OS Hepatitis C virus.

XX WO9813482-A1.

XX 02-APR-1998.

XX 23-SEP-1997; 97WO-US017029.

XX 27-SEP-1996; 96US-0027274P.

PR 12-DEC-1996; 96GB-00025802.

XX (MERI) MERCK & CO INC.

XX Sardana VV, Blue JT;

XX WPI; 1998-230696/20.

XX Detergent-free hepatitis C virus protease NS3 - useful for screening for
PT specific inhibitors and for diagnosing infection, is more active and
PT stable than known enzyme preparations.

XX Disclosure; Page 16-17; 24pp; English.

XX The present sequence represents the nonstructural proteinA (NSA) of
CC Hepatitis C virus (HCV). The specification describes a stable, detergent-
CC free NS3 protease. The specification describes a screening assay for
CC compounds that inhibit NS3. The assay comprises incubating NS3 with the
CC test compound and detecting any inhibition of protease activity that
CC occurs. NS3 can be used to screen for inhibitory compounds and for
CC diagnosis of HCV infection
XX

SQ Sequence 54 AA;

Query Match 10.2%; Score 12; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 7 GGVLAALAAAYCL 18

RESULT 300

AAW17898
ID AAW17898 standard; peptide; 54 AA.
XX
AC AAW17898;

XX 07-SEP-1999 (first entry)

XX Native HCV NS4A peptide.

XX HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.
XX

OS Hepatitis C virus.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

PR 28-JUL-1998; 98US-0094331P.

XX (SCHE) SCHERING CORP.

XX Malcolm BA, Taremi SS, Weber PC, Yao N;

XX WPI; 1999-385385/32.

PT New hepatitis C virus covalent complexes.

XX Disclosure; Page 110; 211pp; English.

XX The present invention describes a covalent hepatitis C virus (HCV) NS4A-
CC NS3 complex comprising a central hydrophobic domain of native HCV NS4A
CC peptide, a linker, and an HCV NS3 serine protease domain, where the
CC hydrophobic domain of native HCV NS4A peptide is tethered by the linker
CC to the amino terminus of the HCV NS3 protease domain. The covalent NS4A-
CC NS3 complexes are useful for structural determination and determination
CC of mode of binding of HCV inhibitors by NMR spectroscopy. They can also
CC be used for detecting inhibitors of the protease activity, the helicase
CC activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes
CC are more soluble, stable and active than the non-covalent protease-
CC peptide complexes previously available. The present sequence represents
CC the native NS4A peptide
XX

SQ Sequence 54 AA;

Query Match 10.2%; Score 12; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 7 GGVLAALAAAYCL 18

RESULT 301

ADP04992
ID ADP04992 standard; protein; 54 AA.

XX ADP04992;

XX 26-AUG-2004 (first entry)

XX Hepatitis C virus NS4A protein, SEQ ID NO:3.

XX HCV; non-structural protein 4A; NS4A; expression enhancer;
KW transcriptional enhancer; immunogenicity enhancer; antigen;
KW genetic vaccine.
XX

OS Hepatitis C virus.

XX WO2004048403-A2.
XX PD 10-JUN-2004.
XX PF 25-NOV-2003; 2003WO-IB006488.
XX PR 26-NOV-2002; 2002US-0430009P.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M;
XX PI WPI; 2004-441149/41.
XX DR N-P5DB; ADP04993.
XX DR Increasing the expression of a nucleic acid in a cell comprises providing
XX PT a first nucleic acid encoding a hepatitis C virus (HCV) non-structural
XX PT protein 4A (NS4A) or its functional portion.
XX PS Claim 10; SEQ ID NO 2; 43pp; English.
XX CC The invention relates to a method of increasing the expression of a
XX CC desired nucleic acid in a cell by associating it with a nucleic acid
XX CC encoding the hepatitis C virus (HCV) non-structural protein 4A (NS4A) or
XX CC a functional portion thereof. The desired nucleic acid to be expressed is
XX CC preferably one encoding the HCV non-structural protein 3 (NS3), and the
XX CC desired nucleic acid and the NS4A gene acts as an enhancer that increases
XX CC both the transcription and also the immunogenicity of an associated
XX CC nucleic acid such as the NS3 gene. The invention also relates to a method
XX CC of increasing immunogenicity of an antigen (preferably the HCV NS3
XX CC protein) in a mammal by associating the antigen-encoding nucleic acid
XX CC with an NS4A nucleic acid, again either on the same or on different
XX CC constructs. The methods of the invention are useful for increasing the
XX CC expression and immunogenicity of a nucleic acid of interest and have
XX CC application in the production of genetic vaccines. The present sequence
XX CC represents the HCV NS4A protein, a nucleic acid encoding all or part of
XX CC which may be used as an enhancer in the method of the invention.
XX SQ Sequence 54 AA;
Query Match 10.2%; Score 12; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 18 GGVLAALAAAYCL 29
Db 7 GGVLAALAAAYCL 18
RESULT 302
ADP05000
ID ADP05000 standard; protein; 54 AA.
XX AC ADP05000;
XX DT 26-AUG-2004 (first entry)
XX DE Hepatitis C virus mutant NS4A protein S1P.
XX KW HCV; non-structural protein 4A; NS4A; expression enhancer;
XX KW transcriptional enhancer; immunogenicity enhancer; antigen;
XX KW genetic vaccine; mutant; mutein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1
XX FT /note= "Pro replaces wild-type Ser"
XX FT
XX PN WO2004048403-A2.

XX 10-JUN-2004.
XX PD 25-NOV-2003; 2003WO-IB006488.
XX PF 26-NOV-2002; 2002US-0430009P.
XX PR (TRIP-) TRIPEP AB.
XX PA Sallberg M;
XX PI WPI; 2004-441149/41.
XX DR Increasing the expression of a nucleic acid in a cell comprises providing
XX PT a first nucleic acid encoding a hepatitis C virus (HCV) non-structural
XX PT protein 4A (NS4A) or its functional portion.
XX PS Example 1; Page; 43pp; English.
XX CC The invention relates to a method of increasing the expression of a
XX CC desired nucleic acid in a cell by associating it with a nucleic acid
XX CC encoding the hepatitis C virus (HCV) non-structural protein 4A (NS4A) or
XX CC a functional portion thereof. The desired nucleic acid to be expressed is
XX CC preferably one encoding the HCV non-structural protein 3 (NS3), and the
XX CC desired nucleic acid and the NS4A gene acts as an enhancer that increases
XX CC both the transcription and also the immunogenicity of an associated
XX CC nucleic acid such as the NS3 gene. The invention also relates to a method
XX CC of increasing immunogenicity of an antigen (preferably the HCV NS3
XX CC protein) in a mammal by associating the antigen-encoding nucleic acid
XX CC with an NS4A nucleic acid, again either on the same or on different
XX CC constructs. The methods of the invention are useful for increasing the
XX CC expression and immunogenicity of a nucleic acid of interest and have
XX CC application in the production of genetic vaccines. The present sequence
XX CC represents a mutant HCV NS4A protein, a nucleic acid encoding which was
XX CC used in an example of the invention. Note: The present sequence is not
XX CC shown in the patent, but is derived from the wild-type HCV NS4A sequence
XX CC (ADP04992) and the information given on page 8.
XX SQ Sequence 54 AA;
Query Match 10.2%; Score 12; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 18 GGVLAALAAAYCL 29
Db 7 GGVLAALAAAYCL 18
RESULT 303
AAR22439
ID AAR22439 standard; protein; 114 AA.
XX AC AAR22439;
XX DT 27-AUG-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 20-AUG-1992 (first entry)
XX DE Non-A non-B hepatitis virus antigen peptide.
XX KW NANBH; antibody detection; vaccine.
XX OS Non-A.
XX OS non-B hepatitis virus.
XX PN JP04066596-A.
XX PD 02-MAR-1992.
XX PF 03-JUL-1990; 90JP-00176669.
XX XX

```
PR 03-JUL-1990; 90JP-00176669.
XX
XX PA (KAGA ) KAGAKU OYOBI KESSEI RYOHO.
XX WPI; 1992-120901/15.
DR N-PSDB; AAQ23456.
XX
XX Non-A non-B hepatitis virus antigen peptide - for detection of antibodies
PT and for use in vaccines.
XX
XX Claim 9; Page 2; 11pp; Japanese.
XX
XX The sequence is that of a non-A non-B hepatitis (NANBH) virus antigen
CC which can be used in the detection of NANBH-related antibodies. It is
CC also useful in the prepn. of vaccines for preventing NANBH infection.
CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to
CC correct OS field.)
XX
XX Sequence 114 AA;
SQ
Query Match 10.2%; Score 12; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 56 GGVLAALAAAYCL 67
RESULT 304
AAR23883
ID AAR23883 standard; protein; 124 AA.
XX
XX AC AAR23883;
XX
XX 26-NOV-1992 (first entry)
XX NANBH peptide B.
XX
XX Beta-gal; expression vector; antigen; antibody.
XX
XX Non-A.
XX non-B hepatitis virus.
XX
XX JP04121193-A.
XX
XX 22-APR-1992.
XX
XX 07-SEP-1990; 90JP-00238312.
XX
XX 07-SEP-1990; 90JP-00238312.
XX (KAGA ) KAGAKU OYOBI KESSEI RYOHO.
XX
XX WPI; 1992-188068/23.
DR N-PSDB; AAQ24561.
XX
XX New non-A, non-B hepatitis virus fused peptide - useful as antibody assay
PT reagent for NANBH as it retains both antigenicities and increases
PT detection percentage.
XX
XX Disclosure; Fig 4; 11pp; Japanese.
XX
XX The sequences given in AAR23882-83 are two peptides both of which are
CC encoded by the non-A, non-B hepatitis virus (NANBH). These peptides can
CC be fused to produce a novel peptide (see also AAR23884) which can be
CC linked to a carrier molecule, pref. beta-gal. The DNA encoding the novel
CC peptide can be integrated into an expression vector and this can be use
CC to transform a host cell which can be cultured to express the peptide.
CC This novel peptide can be used as an antigen in an assay method to detect
CC NANBH antibody. The novel peptide:beta-gal fusion retains the
CC antigenicity of both components and can be used in the detection of NANBH
XX
```

```
SQ Sequence 124 AA;
Query Match 10.2%; Score 12; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 56 GGVLAALAAAYCL 67
RESULT 305
AAR25113
ID AAR25113 standard; protein; 134 AA.
XX
XX AC AAR25113;
XX
XX 27-AUG-2003 (revised)
XX 25-MAR-2003 (revised)
DT 07-DEC-1992 (first entry)
XX
XX Non-A, Non-B Hepatitis Virus antigen #4.
XX
XX Antigen S29; NANBH; Hepatitis C; HCV; T064; T069; T06A; ELISA.
XX
XX Non-A.
XX non-B hepatitis virus.
XX
XX WO9209634-A1.
XX
XX 11-JUN-1992.
XX
XX 29-NOV-1991; 91WO-JP001662.
XX
XX 29-NOV-1990; 90JP-00325434.
XX 29-NOV-1990; 90JP-00325435.
XX 16-JAN-1991; 91JP-00070231.
XX 19-APR-1991; 91JP-00179074.
XX 07-JUN-1991; 91JP-00232590.
XX (TORA ) TORAY IND INC.
XX
XX Arima T, Sato A, Ida N, Kazami J;
XX
XX WPI; 1992-217026/26.
DR N-PSDB; AAQ25743.
XX
XX New non-A non-B hepatitis virus antigen proteins - for highly specific
PT detection of hepatitis.
XX
XX Claim 1; Page 32-33; 80pp; Japanese.
XX
XX This sequence is one of 12 claimed antigen sequences specific to NANBH
CC virus. The antigens can be used singly or in combination in an ELISA
CC diagnosis of hepatitis. See AAR24946 and AAR25110-R35121. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
CC field.)
XX
XX Sequence 134 AA;
Query Match 10.2%; Score 12; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 0.0039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 12 GGVLAALAAAYCL 23
RESULT 306
AAR52737
ID AAR52737 standard; protein; 194 AA.
XX
```

AC AAR52737;
XX
DT 31-JAN-1995 (first entry)
XX
DE HCV antigen.
XX
KW Hepatitis C virus; HCV; antigen; diagnosis; reagent; agglutination.
XX
OS Synthetic.
XX
PN JP06102273-A.
XX
PD 15-APR-1994.
XX
PF 18-SEP-1992; 92JP-00250027.
XX
PR 18-SEP-1992; 92JP-00250027.
XX
PA (TOKU) TOKUYAMA SODA KK.
XX
DR WPI; 1994-161280/20.
DR N-PSDB; AAQ62690.
XX
PT Immunological agglutination reagent for the diagnosis of hepatitis C -
PT comprising hepatitis C virus antigen polypeptide.
XX
PS Claim 2-3; Page 17-18; 18pp; Japanese.
XX
CC An new immunological agglutination reaction reagent for the diagnosis of
CC hepatitis C uses a HCV antigen polypeptide subjected to heat-treatment.
CC The HCV antigen active polypeptide contains one of the polypeptides given
CC in AAR52735-38
XX
SQ Sequence 194 AA;

Query Match 10.2%; Score 12; DB 2; Length 194;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 60 GGVLAALAAAYCL 71

RESULT 307
AAR25863
ID AAR25863 standard; protein; 195 AA.
AC AAR25863;
XX
DT 21-JAN-1993 (first entry)
XX
DE HCV polypeptide 10.
XX
KW Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
XX
OS Hepatitis C virus.
XX
PN JP04179482-A.
XX
PD 26-JUN-1992.
XX
PF 11-NOV-1990; 90JP-00304417.
XX
PR 11-NOV-1990; 90JP-00304417.
XX
PA (TOKU) TOKUYAMA SODA KK.
XX
DR WPI; 1992-263663/32.
DR N-PSDB; AAQ26990.
XX
PT Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
PT diagnosis of hepatitis C virus infection.

XX
PS Claim 1; Page 3-4; 66pp; Japanese.
XX
CC The sequences given in AAR25854-74 are hepatitis C virus proteins. The
CC genes encoding these proteins can each be used to prepare recombinant
CC vectors by ligating the gene of interest in to a vector to be expressed
CC in E. coli. These polypeptides are useful as diagnostic reagents for type
CC C hepatitis and they may be produced efficiently by recombinant methods
XX
SQ Sequence 195 AA;

Query Match 10.2%; Score 12; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 61 GGVLAALAAAYCL 72

RESULT 308
AAW41741
ID AAW41741 standard; protein; 195 AA.
XX
AC AAW41741;
XX
DT 22-MAY-1998 (first entry)
XX
DE Hepatitis C virus antigen.
XX
KW non-A non-B hepatitis virus; NANBH; hepatitis C virus; HCV; antigen;
KW diagnosis; detection.
XX
OS Hepatitis virus.
XX
PN JP05176774-A.
XX
PD 20-JUL-1993.
XX
PF 18-DEC-1991; 91JP-00354708.
XX
PR 18-DEC-1990; 90JP-00412020.
XX
PA (SHIM/) SHIMOTONO K.
PA (GREC) GREEN CROSS CORP.
XX
DR WPI; 1993-260858/33.
XX
PT Protein contg. non-A non-B hepatitis antigen fragment - prepd. by
PT culturing transformants transformed by vector contg. base sequence coding
PT specified aminoacid sequences, used for detecting hepatitis.
XX
PS Claim 1; Fig 7; 53pp; Japanese.
XX
CC The present sequence is a non-A non-B hepatitis virus (NANBH) or
CC hepatitis C virus (HCV) antigen, useful for diagnosis or detection
XX
SQ Sequence 195 AA;

Query Match 10.2%; Score 12; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 61 GGVLAALAAAYCL 72

RESULT 309
AAR29846
ID AAR29846 standard; protein; 200 AA.
XX
AC AAR29846;

```

XX 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX
XX HCV NS2-NS4 peptide N13-1.
DB
XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcripase; cDNA; primer; allele.
XX
XX Hepatitis C virus.
OS
XX EP518313-A2.
PN
XX 16-DEC-1992.
PD
XX 11-JUN-1992; 92EP-00109812.
XX
XX 11-JUN-1991; 91JP-00139268.
XX
XX 12-JUL-1991; 91JP-00172794.
PR
XX 07-OCT-1991; 91JP-00287008.
PR
XX 16-DEC-1991; 91JP-00332329.
PR
XX 20-APR-1992; 92JP-00099957.
PR
XX (MITU ) MITSUBISHI KASEI CORP.
PA
XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
PI WPI; 1992-417213/51.
XX
XX N-PSDB; AAQ32477.
DR
XX
XX New hepatitis C virus gene and its encoded protein - used for diagnosing
FT and vaccinating against hepatitis C virus infections.
PT
XX
XX Disclosure; Page 125-26; 305pp; English.
PS
XX
XX The sequences given in AAR29559-60 and AAR29843-51 were encoded
CC by clones which encode the NS2-NS4 regions of the Hepatitis C virus (HCV)
CC gene of the invention. These sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The NS2-NS4 RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32553-64. The cDNA sequences were then
CC amplified using primer pairs. The cDNA sequences isolated represent
CC different alleles of the same region of the HCV gene. Sequence
CC comparisons of these clones showed that it is possible for a patient to
CC carry more than one HCV strain at one time. See also AAQ32436. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 200 AA;
SQ
Query Match 10.2%; Score 12; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 101 GGVLAALAAAYCL 112
RESULT 310
AAR25885
ID AAR25885 standard; protein; 222 AA.
XX
XX AAR25885;
AC
XX 09-SEP-2004 (revised)
DT 21-JAN-1993 (first entry)
XX
XX HK10.
DE
XX Recombinant vector; E. coli; diagnostic; reagent; type C hepatitis.
KW
XX Hepatitis C virus.
OS
XX Unidentified.

```

```

XX Key Location/Qualifiers
FH Protein 23..217
FT /note= "Sequence AAR25863"
XX
XX JP04179482-A.
PN
XX 26-JUN-1992.
XX
XX 11-NOV-1990; 90JP-00304417.
XX
XX 11-NOV-1990; 90JP-00304417.
PR
XX (TOKU ) TOKUYAMA SODA KK.
PA
XX WPI; 1992-263663/32.
XX
XX N-PSDB; AAQ27012.
DR
XX Hepatitis C virus antigen expressed as recombinant in E.coli - useful for
FT diagnosis of hepatitis C virus infection.
XX
XX Disclosure; Fig 11; 66pp; Japanese.
PS
XX
XX The sequences given in AAR25876-95 are encoded by the claimed hepatitis C
CC virus genes of the invention which have been inserted into an E. coli
CC vector. These polypeptides are useful as diagnostic reagents for type C
CC hepatitis and they may be produced efficiently by recombinant DNA
CC techniques
CC
XX Revised record issued on 09-SEP-2004 : Correction to feature table key
CC
XX Sequence 222 AA;
SQ
Query Match 10.2%; Score 12; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.006;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 83 GGVLAALAAAYCL 94
RESULT 311
AAR23884
ID AAR23884 standard; protein; 272 AA.
XX
XX AAR23884;
AC
XX 26-NOV-1992 (first entry)
DT
XX NANBH fusion peptide.
DE
XX Beta-gal; expression vector; antigen; antibody.
KW
XX Non-A.
OS
XX non-B hepatitis virus.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..143
FT /label= peptide_A
FT Region 144..145
FT /label= linker
FT Peptide 146..272
FT /label= peptide_B
XX
XX JP04121193-A.
PN
XX 22-APR-1992.
XX
XX 07-SEP-1990; 90JP-00238312.
XX
XX 07-SEP-1990; 90JP-00238312.
XX

```


PA (KAGA) KAGAKU OYOBI KESSEI RYOHO.
XX WPI; 1992-188068/23.
DR N-PSDB; AAQ24562.
XX
PT New non-A, non-B hepatitis virus fused peptide - useful as antibody assay
PT reagent for NANBH as it retains both antigenicities and increases
PT detection percentage.
XX
PS Claim 6; Page 2; 11pp; Japanese.
XX
CC The sequence given is a novel peptide formed by the fusion of two
CC peptides both of which are encoded by the non-A, non-B hepatitis virus
CC (NANBH) (see also AAR23882-3). This novel peptide can be linked to a
CC carrier molecule, pref. beta-gal. The DNA encoding the novel peptide can
CC be integrated into an expression vector and this can be used to transform
CC a host cell which can be cultured to express the peptide. This novel
CC peptide can be used as an antigen in an assay method to detect NANBH
CC antibody. The novel peptide:beta-gal fusion retains the antigenicity of
CC both components and can be used in the detection of NANBH
XX
SQ Sequence 272 AA;
Query Match 10.2%; Score 12; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 204 GGVLAALAAAYCL 215
RESULT 312
AAR98350
ID AAR98350 standard; peptide; 293 AA.
XX
AC AAR98350;
XX
DT 21-NOV-1996 (first entry)
XX
DE HCV NS4 derived antigen.
XX
XX Antigen; NS4; hepatitis C virus; HCV; antibody; bird; egg yolk;
KW detection; NS3; blood; diagnosis.
XX
OS Hepatitis C virus.
XX
XX JP08127598-A.
PN
PD 21-MAY-1996.
XX
PF 28-OCT-1994; 94JP-00264808.
XX
PR 28-OCT-1994; 94JP-00264808.
XX
PA (SHIE) SHINETSU CHEM IND CO LTD.
PA (MITU) MITSUBISHI CHEM CORP.
XX
DR WPI; 1996-295556/30.
XX
XX Anti-HCV NS3 and NS4 protein antibodies - produced by immunising a bird
PT and isolating the antibodies from egg yolk.
PT
PS Claim 2; Page 7-8; 9pp; English.
XX
CC The sequences given in AAR98349-50 represent antigenic peptides which
CC were derived from the NS3 and NS4 proteins of hepatitis C virus (HCV),
CC respectively. These peptides were used in the preparation of novel
CC antibodies. The antibodies were prepared by immunising a bird with one of
CC the peptides and isolating the resulting antibody from the egg yolk of
CC eggs laid by the immunised bird. The antibodies may be used in the
CC detection of HCV antigens in the blood and are useful in the development
CC of diagnostic agents or drugs for hepatitis C. The antibody is highly

CC specific and as the peptide is injected without a carrier protein, only
CC anti-NS3 or anti-NS4 antibodies are produced
XX
SQ Sequence 293 AA;
Query Match 10.2%; Score 12; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 101 GGVLAALAAAYCL 112
RESULT 313
AAR29907
ID AAR29907 standard; protein; 320 AA.
XX
AC AAR29907;
XX
DT 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX
XX HCV antigen peptide O15-1.
DE
XX Clone; Hepatitis C Virus; HCV; core-envelope; NS1(gp70); NS2-NS4;
KW NS4-NS5; region; diagnostic method; antibody; suppress; control;
KW proteolytic; process; precursor; polypeptide.
XX
OS Hepatitis C virus.
XX
PN EP518313-A2.
XX
PD 16-DEC-1992.
XX
PF 11-JUN-1992; 92EP-00109812.
XX
PR 11-JUN-1991; 91JP-00139268.
PR 12-JUN-1991; 91JP-00172794.
PR 07-OCT-1991; 91JP-00287008.
PR 16-DEC-1991; 91JP-00332329.
PR 20-APR-1992; 92JP-00099957.
XX
PA (MITU) MITSUBISHI KASEI CORP.
XX
PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
XX WPI; 1992-417213/51.
DR N-PSDB; AAQ32538.
XX
PT New hepatitis C virus gene and its encoded protein - used for diagnosing
PT and vaccinating against hepatitis C virus infections.
XX
PS Disclosure; Page 275-276; 305pp; English.
XX
XX The sequences given in AAR29528 and AAR29907-08 are encoded by various
CC clones of the full length Hepatitis C Virus (HCV) gene of the invention.
CC These HCV peptides may be useful in the development of a diagnostic
CC method which is more accurate and effective than conventional ones, in
CC the detection of antibodies raised against a wide range of HCVs which
CC have been hardly detected before. The complete gene may be used in an
CC vitro screening system for a substance capable of specifically suppressing
CC or controlling a proteolytic processing of a precursor polypeptide of
CC HCV. See also AAQ32436. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 320 AA;
Query Match 10.2%; Score 12; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 0.0082;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
XXXXXXXXXXXX

PN WO9317110-A2.
 XX 02-SEP-1993.
 XX 19-FEB-1993; 93WO-GB000345.
 XX 21-FEB-1992; 92GB-0003803.
 XX (WELL) WELLCOME FOUND LTD.
 PA Parker D, Rodgers BC;
 XX
 XX WPI: 1993-288415/36.
 DR N-PSDB; AAQ46198.
 XX
 PT New recombinant polypeptide for diagnosing hepatitis C - contains three
 PT distinct antigens from different viral regions, also useful in protective
 PT vaccines.
 XX
 PS Example 1; Page 64-66; 99pp; English.
 XX
 CC The NS4 region from the 3' region of the PT-NANBH genome (AAQ46195) is
 CC amplified by PCR using primers D224 and D226 (AAQ46196-97) and the
 CC fragment (AAQ46198) is cloned into a vector and expressed in infected
 CC insect cells. The recombinant virus (BHC-19) was able to express the NS4
 CC specific recombinant protein at low levels in the infected insect cells.
 CC If at least three different PT-NANBH antigens are used to screen for PT-
 CC NANBH, the screening is much more sensitive as compared to the use of
 CC only two PT-NANBH antigens. Pref. antigens are described in AAQ46192-94.
 CC Two new antigenic regions of the PT-NANBH genome are given in AAQ46198-
 CC 99. AAQ46202 describes an improved PT-NANBH recombinant polypeptide.
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX
 XX Sequence 373 AA;

Query Match 10.2%; Score 12; DB 2; Length 373;
 Best Local Similarity 100.0%; Pred. No. 0.0093;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 99 GGVLAALAAAYCL 110

RESULT 317
 AAR29866
 ID AAR29866 standard; protein; 477 AA.
 XX
 AC AAR29866;

DT 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)
 XX
 DE HCV NS2-NS4 peptide N16N15B-1.

XX
 KW Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
 KW transcriptase; cDNA; primer; allele.

OS Hepatitis C virus.

XX
 XX
 FN EP518313-A2.

XX
 PD 16-DEC-1992.

XX
 PF 11-JUN-1992; 92EP-00109812.

XX
 PR 11-JUN-1991; 91JP-00139268.

PR 12-JUL-1991; 91JP-00172794.

PR 07-OCT-1991; 91JP-00287008.

PR 16-DEC-1991; 91JP-00332329.

PR 20-APR-1992; 92JP-00099957.

XX

PA (MITU) MITSUBISHI KASEI CORP.

XX
 PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX
 DR WPI: 1992-417213/51.

XX
 DR N-PSDB; AAQ32497.

PT New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.

XX
 PS Disclosure; Page 172-75; 305pp; English.

XX
 CC The sequences given in AAR29852-70 are encoded by various clones which
 CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C
 CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and
 CC AAR29843-51). These RNA sequences were isolated from the serum of a
 CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
 CC converted into cDNA using transcriptase in the presence of one of the
 CC primer sequences given in AAQ32578-79. The sequences were then amplified
 CC using primer pairs. The cDNA sequences isolated represent different
 CC alleles of the same region of the HCV gene. Sequence comparisons of these
 CC clones showed that it is possible for a patient to carry more than one
 CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX

SQ Sequence 477 AA;

Query Match 10.2%; Score 12; DB 2; Length 477;
 Best Local Similarity 100.0%; Pred. No. 0.011;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 222 GGVLAALAAAYCL 233

RESULT 318
 AAR29867
 ID AAR29867 standard; protein; 477 AA.
 XX
 AC AAR29867;

XX
 DT 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)

XX
 DE HCV NS2-NS4 peptide N16N15-1.

XX
 KW Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
 KW transcriptase; cDNA; primer; allele.

XX
 OS Hepatitis C virus.

XX
 FN EP518313-A2.

XX
 PD 16-DEC-1992.

XX
 PF 11-JUN-1992; 92EP-00109812.

XX
 PR 11-JUN-1991; 91JP-00139268.

PR 12-JUL-1991; 91JP-00172794.

PR 07-OCT-1991; 91JP-00287008.

PR 16-DEC-1991; 91JP-00332329.

PR 20-APR-1992; 92JP-00099957.

XX
 PA (MITU) MITSUBISHI KASEI CORP.

XX
 PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX
 DR WPI: 1992-417213/51.

XX
 DR N-PSDB; AAQ32498.

PT New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 175-77; 305pp; English.

XX The sequences given in AAR29852-70 are encoded by various clones which

CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C

CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and

CC AAR29843-51). These RNA sequences were isolated from the serum of a

CC patient suffering from hepatitis C (HC). The isolated RNA sequences were

CC converted into cDNA using transcriptase in the presence of one of the

CC primer sequences given in AAR29852-79. The sequences were then amplified

CC using primer pairs. The cDNA sequences isolated represent different

CC alleles of the same region of the HCV gene. Sequence comparisons of these

CC clones showed that it is possible for a patient to carry more than one

CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to

CC correct PN field.)

XX Sequence 477 AA;

Query Match 10.2%; Score 12; DB 2; Length 477;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 222 GGVLAALAAAYCL 233

RESULT 319

AAR29865

ID AAR29865 standard; protein; 477 AA.

AC AAR29865;

XX 25-MAR-2003 (revised)

DT 26-APR-1993 (first entry)

DE HCV NS2-NS4 peptide N16N15A-1.

XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;

KW transcriptase; cDNA; primer; allele.

XX Hepatitis C virus.

XX EP518313-A2.

XX 16-DEC-1992.

XX 11-JUN-1992; 92EP-00109812.

XX 11-JUN-1991; 91JP-00139268.

XX 12-JUL-1991; 91JP-00172794.

XX 07-OCT-1991; 91JP-00287008.

XX 16-DEC-1991; 91JP-00332329.

XX 20-APR-1992; 92JP-00099957.

XX (MITU) MITSUBISHI KASEI CORP.

XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX WPI; 1992-417213/51.

XX N-PSDB; AAQ32496.

XX New hepatitis C virus gene and its encoded protein - used for diagnosing

PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 170-72; 305pp; English.

XX The sequences given in AAR29852-70 are encoded by various clones which

CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C

CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and

CC AAR29843-51). These RNA sequences were isolated from the serum of a

CC patient suffering from hepatitis C (HC). The isolated RNA sequences were

CC converted into cDNA using transcriptase in the presence of one of the

CC primer sequences given in AAR29852-79. The sequences were then amplified

CC using primer pairs. The cDNA sequences isolated represent different

CC alleles of the same region of the HCV gene. Sequence comparisons of these

CC clones showed that it is possible for a patient to carry more than one

CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to

CC correct PN field.)

XX Sequence 477 AA;

Query Match 10.2%; Score 12; DB 2; Length 477;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 222 GGVLAALAAAYCL 233

RESULT 320

ADZ99621

ID ADZ99621 standard; protein; 685 AA.

XX ADZ99621;

XX 14-JUL-2005 (first entry)

DE Hepatitis C virus (HCV) NS3/4A serine protease mutant protein 2.

XX protein engineering; antiinflammatory; hepatotropic; virucide;

KW hepatitis C virus infection; NS3; NS4A; serine protease; enzyme; muten.

XX Hepatitis C virus.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 168 /label= Asp, Glu

FT /note= "Wild-type Asp may be substituted"

XX WO2005042570-A1.

XX 12-MAY-2005.

XX 27-OCT-2004; 2004WO-US035839.

XX 27-OCT-2003; 2003US-0514740P.

XX 26-NOV-2003; 2003US-0525222P.

XX 13-APR-2004; 2004US-0561662P.

XX (VERT-) VERTEX PHARM INC.

XX Lin C, Lin K;

XX WPI; 2005-366577/37.

XX Drug resistant mutants of hepatitis C virus (HCV) NS3/4A protease, useful

PT in assessing HCV infections in patients, and in screening for new

PT antiviral drugs.

XX Claim 17; Page; 113pp; English.

XX The invention relates to a novel isolated Hepatitis C virus (HCV)

CC polynucleotide encoding an HCV NS3/4A protease or its biologically active

CC analog, where the codon that corresponds to codon 156 of the wild-type

CC polynucleotide and/or the codon that corresponds to codon 168 of the wild

CC -type polynucleotide is mutated so that it does not encode an alanine at

CC 156 and/or aspartic acid at 168. The HCV NS protein 3 (NS3) contains a

CC serine protease activity that processes the viral polyprotein to generate

CC the majority of the viral enzymes, and is essential for viral replication

CC and infectivity. NS4A is an associated cofactor of NS3. The

CC polynucleotide of the invention demonstrates antiinflammatory,

CC hepatotropic and virucide activities and may be useful in assessing

CC Hepatitis C virus (HCV) infections in patients, as well as in screening

CC for new therapeutics targeting the HCV NS3/4A protease. Compounds
CC identified by the methods may be useful for treating HCV infections and
CC for eliminating or reducing HCV contamination of biological samples and
CC medical and laboratory equipment. The current sequence is that of the HCV
CC NS3/4A serine protease mutant protein 2 of the invention. The current
CC sequence is not shown within the specification per se but was created by
CC the indexer using information in claim 17.

XX SQ Sequence 685 AA;
Query Match 10.2%; Score 12; DB 9; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 638 GGVLAALAAAYCL 649

RESULT 321
ADZ99616
ID ADZ99616 standard; protein; 685 AA.

AC ADZ99616;

XX 14-JUL-2005 (first entry)

XX Hepatitis C virus (HCV) NS3/4A serine protease wild-type protein.

XX protein engineering; antiinflammatory; hepatotropic; virucide;

KW hepatitis C virus infection; NS3; NS4A; serine protease; enzyme.

XX Hepatitis C virus.

OS WO2005042570-A1.

PN 12-MAY-2005.

XX 27-OCT-2004; 2004WO-US035839.

XX 27-OCT-2003; 2003US-0514740P.

PR 26-NOV-2003; 2003US-0525222P.

PR 13-APR-2004; 2004US-0561662P.

XX (VERT-) VERTEX PHARM INC.

XX Lin C, Lin K;

XX WPI; 2005-366577/37.

DR N-PSDB; ADZ99615.

XX Drug resistant mutants of hepatitis C virus (HCV) NS3/4A protease, useful
PT in assessing HCV infections in patients, and in screening for new
PT antiviral drugs.

PS Claim 20; SEQ ID NO 2; 113pp; English.

XX The invention relates to a novel isolated Hepatitis C virus (HCV)
CC polynucleotide encoding an HCV NS3/4A protease or its biologically active
CC analog, where the codon that corresponds to codon 156 of the wild-type
CC polynucleotide and/or the codon that corresponds to codon 168 of the wild
CC -type polynucleotide is mutated so that it does not encode an alanine at
CC 156 and/or aspartic acid at 168. The HCV NS protein 3 (NS3) contains a
CC serine protease activity that processes the viral polyprotein to generate
CC the majority of the viral enzymes, and is essential for viral replication
CC and infectivity. NS4A is an associated cofactor of NS3. The
CC polynucleotide of the invention demonstrates antiinflammatory,

CC hepatotropic and virucide activities and may be useful in assessing
CC Hepatitis C virus (HCV) infections in patients, as well as in screening
CC for new therapeutics targeting the HCV NS3/4A protease. Compounds
CC identified by the methods may be useful for treating HCV infections and
CC for eliminating or reducing HCV contamination of biological samples and
CC medical and laboratory equipment. The current sequence is that of the HCV

CC NS3/4A serine protease wild-type protein of the invention.

XX SQ Sequence 685 AA;
Query Match 10.2%; Score 12; DB 9; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.015; 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 638 GGVLAALAAAYCL 649

RESULT 322
ADZ99623
ID ADZ99623 standard; protein; 685 AA.

AC ADZ99623;

XX 14-JUL-2005 (first entry)

XX Hepatitis C virus (HCV) NS3/4A serine protease mutant protein 4.

XX protein engineering; antiinflammatory; hepatotropic; virucide;

KW hepatitis C virus infection; NS3; NS4A; serine protease; enzyme; mutein.

XX Hepatitis C virus.

OS Synthetic.

XX WO2005042570-A1.

XX 12-MAY-2005.

XX 27-OCT-2004; 2004WO-US035839.

XX 27-OCT-2003; 2003US-0514740P.

PR 26-NOV-2003; 2003US-0525222P.

PR 13-APR-2004; 2004US-0561662P.

XX (VERT-) VERTEX PHARM INC.

XX Lin C, Lin K;

XX WPI; 2005-366577/37.

XX Drug resistant mutants of hepatitis C virus (HCV) NS3/4A protease, useful
PT in assessing HCV infections in patients, and in screening for new
PT antiviral drugs.

PS Claim 19; Page; 113pp; English.

XX The invention relates to a novel isolated Hepatitis C virus (HCV)
CC polynucleotide encoding an HCV NS3/4A protease or its biologically active
CC analog, where the codon that corresponds to codon 156 of the wild-type
CC polynucleotide and/or the codon that corresponds to codon 168 of the wild
CC -type polynucleotide is mutated so that it does not encode an alanine at
CC 156 and/or aspartic acid at 168. The HCV NS protein 3 (NS3) contains a
CC serine protease activity that processes the viral polyprotein to generate
CC the majority of the viral enzymes, and is essential for viral replication
CC and infectivity. NS4A is an associated cofactor of NS3. The
CC polynucleotide of the invention demonstrates antiinflammatory,
CC hepatotropic and virucide activities and may be useful in assessing
CC Hepatitis C virus (HCV) infections in patients, as well as in screening
CC for new therapeutics targeting the HCV NS3/4A protease. Compounds
CC identified by the methods may be useful for treating HCV infections and
CC for eliminating or reducing HCV contamination of biological samples and
CC medical and laboratory equipment. The current sequence is that of the HCV
CC NS3/4A serine protease mutant protein 4 of the invention. The current
CC sequence is not shown within the specification per se but was created by
CC the indexer using information in claim 19.

XX SQ Sequence 685 AA;

Query Match 10.2%; Score 12; DB 9; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 638 GGVLAAALAAAYCL 649

RESULT 323
ADZ99622
ID ADZ99622 standard; protein; 685 AA.

XX ADZ99622;
XX 14-JUL-2005 (first entry)
XX Hepatitis C virus (HCV) NS3/4A serine protease mutant protein 3.
XX protein engineering; antiinflammatory; hepatotropic; virucide;
XX hepatitis C virus infection; NS3; NS4A; serine protease; enzyme; mutein.
XX Hepatitis C virus.
XX Synthetic.

PN WO2005042570-A1.
XX 12-MAY-2005.
XX 27-OCT-2004; 2004WO-US035839.
XX 27-OCT-2003; 2003US-0514740P.
XX 26-NOV-2003; 2003US-0525222P.
XX 13-APR-2004; 2004US-0561662P.

(VERT-) VERTEX PHARM INC.

Lin C, Lin K;

WPI; 2005-366577/37.

Drug resistant mutants of hepatitis C virus (HCV) NS3/4A protease, useful
in assessing HCV infections in patients, and in screening for new
antiviral drugs.

Claim 18; Page; 113pp; English.

The invention relates to a novel isolated Hepatitis C virus (HCV)
polynucleotide encoding an HCV NS3/4A protease or its biologically active
analogs, where the codon that corresponds to codon 156 of the wild-type
polynucleotide and/or the codon that corresponds to codon 168 of the wild
-type polynucleotide is mutated so that it does not encode an alanine at
156 and/or aspartic acid at 168. The HCV NS protein 3 (NS3) contains a
serine protease activity that processes the viral polyprotein to generate
the majority of the viral enzymes, and is essential for viral replication
and infectivity. NS4A is an associated cofactor of NS3. The
polynucleotide of the invention demonstrates antiinflammatory,
hepatotropic and virucide activities and may be useful in assessing
Hepatitis C virus (HCV) infections in patients, as well as in screening
for new therapeutics targeting the HCV NS3/4A protease. Compounds
identified by the methods may be useful for treating HCV infections and
for eliminating or reducing HCV contamination of biological samples and
medical and laboratory equipment. The current sequence is that of the HCV
NS3/4A serine protease mutant protein 3 of the invention. The current
sequence is not shown within the specification per se but was created by
the indexer using information in claim 18.

Sequence 685 AA;

Query Match 10.2%; Score 12; DB 9; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 638 GGVLAAALAAAYCL 649

RESULT 324

ADZ99620
ID ADZ99620 standard; protein; 685 AA.

XX ADZ99620;

XX 14-JUL-2005 (first entry)

XX Hepatitis C virus (HCV) NS3/4A serine protease mutant protein 1.

XX protein engineering; antiinflammatory; hepatotropic; virucide;
XX hepatitis C virus infection; NS3; NS4A; serine protease; enzyme; mutein.

XX Hepatitis C virus.

XX Synthetic.

PN WO2005042570-A1.

XX 12-MAY-2005.

XX 27-OCT-2004; 2004WO-US035839.

XX 27-OCT-2003; 2003US-0514740P.

XX 26-NOV-2003; 2003US-0525222P.

XX 13-APR-2004; 2004US-0561662P.

(VERT-) VERTEX PHARM INC.

Lin C, Lin K;

WPI; 2005-366577/37.

Drug resistant mutants of hepatitis C virus (HCV) NS3/4A protease, useful
in assessing HCV infections in patients, and in screening for new
antiviral drugs.

Claim 16; Page; 113pp; English.

The invention relates to a novel isolated Hepatitis C virus (HCV)
polynucleotide encoding an HCV NS3/4A protease or its biologically active
analogs, where the codon that corresponds to codon 156 of the wild-type
polynucleotide and/or the codon that corresponds to codon 168 of the wild
-type polynucleotide is mutated so that it does not encode an alanine at
156 and/or aspartic acid at 168. The HCV NS protein 3 (NS3) contains a
serine protease activity that processes the viral polyprotein to generate
the majority of the viral enzymes, and is essential for viral replication
and infectivity. NS4A is an associated cofactor of NS3. The
polynucleotide of the invention demonstrates antiinflammatory,
hepatotropic and virucide activities and may be useful in assessing
Hepatitis C virus (HCV) infections in patients, as well as in screening
for new therapeutics targeting the HCV NS3/4A protease. Compounds
identified by the methods may be useful for treating HCV infections and
for eliminating or reducing HCV contamination of biological samples and
medical and laboratory equipment. The current sequence is that of the HCV
NS3/4A serine protease mutant protein 1 of the invention. The current
sequence is not shown within the specification per se but was created by
the indexer using information in claim 16.

Sequence 685 AA;

Query Match 10.2%; Score 12; DB 9; Length 685;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 638 GGVLAAALAAAYCL 649

RESULT 325
 AAR58591
 ID AAR58591 standard; protein; 697 AA.
 XX AC AAR58591;
 XX DT 09-MAY-1995 (first entry)
 XX DE Hepatitis C virus 4725-6817 fragment antigen.
 XX KW Non-A non-B hepatitis virus antigens; NANBH; hepatitis C virus.
 XX OS Hepatitis C virus.
 XX PN JP06225770-A.
 XX PD 16-AUG-1994.
 XX PF 08-JUL-1993; 93JP-00193104.
 XX PR 10-JUL-1992; 92JP-00207391.
 XX PA (TOKR-) ZH TOKYO RINSHO IGAKU SOGO KENKYUSHO.
 XX PA (SANW) SANWA KAGAKU KENKYUSHO CO.
 XX PA (TOFU) TONEN CORP.
 XX PA (KOKU-) KOKUSAI SHIYAKU KK.
 XX WPI; 1994-298800/37.
 XX N-PSDB; AAR70541.
 A nucleic acid fragment coding Non-A Non-B Hepatitis virus antigens - for
 diagnosis of NANBH and detection of HCV.
 Claim 1; Page 14; 22pp; Japanese.
 AAQ70541 is a fragment of hepatitis C virus (HCV) or non-A non-B hepatitis
 virus (NANBH) non-structural coding region four, encompassing base pairs
 4725-6817. It codes for AAR58591 an antigen to a non-structural region of
 the HCV virus, which can be used in the diagnosis of NANBH patients and
 the detection of HCV carriers
 Sequence 697 AA;
 Query Match 10.2%; Score 12; DB 2; Length 697;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCL 29
 DB 206 GGVLAALAAAYCL 217
 RESULT 326
 ADL17782
 ID ADL17782 standard; protein; 697 AA.
 XX AC ADL17782;
 XX DT 06-MAY-2004 (first entry)
 XX DE Hepatitis C virus protein (clone C6-79).
 XX KW non-A non-B type hepatitis virus antigen;
 XX KW hepatitis C virus antibody detection; clone C260-1.
 XX OS Hepatitis C virus.
 XX PN JP2004000150-A.
 XX PD 08-JAN-2004.
 XX PF 24-FEB-2003; 2003JP-00046379.

XX 10-JUL-1992; 92JP-00207391.
 PR 08-JUL-1993; 93JP-00193104.
 XX (KOKU-) KOKUSAI SHIYAKU KK.
 XX WPI; 2004-085213/09.
 DR N-PSDB; ADL17777.
 XX Novel non-A non-B type hepatitis virus nucleic acid fragment useful for
 diagnosing non-A non-B hepatitis patient and non-A non-B type hepatitis
 virus carrier.
 XX Disclosure; Page 30-32; 35pp; Japanese.
 XX This invention relates to a novel isolated nucleic acid fragment encoding
 a non-A non-B type hepatitis virus antigen protein sequence.
 CC Specifically, it refers to a method for diagnosing non-A non-B type
 CC hepatitis patients, as well as for detecting those people who are
 CC carriers. The present invention describes a diagnosis method that uses a
 CC hepatitis C virus antibody detection reagent and subsequent PCR
 CC amplification to identify a clone C260-1 that encodes the non-A non-B
 CC type hepatitis virus antigen sequence. This polypeptide sequence is the
 CC non-A non-B type hepatitis virus protein fragment of the invention. NOTE:
 CC This sequence is given in the sequence listing as an embedded protein and
 CC is not further referred to in the specification.
 XX SQ Sequence 697 AA;
 Query Match 10.2%; Score 12; DB 8; Length 697;
 Best Local Similarity 100.0%; Pred. No. 0.016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCL 29
 DB 206 GGVLAALAAAYCL 217
 RESULT 327
 AAR80044
 ID AAR80044 standard; protein; 767 AA.
 XX AC AAR80044;
 XX DT 25-MAR-1996 (first entry)
 XX DE Hepatitis C virus antigenic sequence.
 XX KW Maltose binding protein; MBP; non-A non-B; HCV; diagnosis; antigen.
 XX OS Hepatitis C virus.
 XX PN JP07198723-A.
 XX PD 01-AUG-1995.
 XX PF 29-DEC-1993; 93JP-00351227.
 XX PR 29-DEC-1993; 93JP-00351227.
 XX PA (JAPG) NIPPON ZEON KK.
 XX WPI; 1995-300583/39.
 DR N-PSDB; AAT04565.
 XX Diagnosis of HCV infection - using a fused protein comprising a HCV
 PT antigen and a carrier protein.
 XX Example 2; Page 7-10; 10pp; Japanese.
 XX AAT04565 encodes AAR80044 a hepatitis C virus (HCV) antigenic sequence.
 CC An antigen derived from the antigenic sequence was combined with a
 CC carrier protein (maltose binding protein) to produce a fusion protein,

CC which can be used for the highly sensitive detection of HCV

XX Sequence 767 AA;

Query Match 10.2%; Score 12; DB 2; Length 767;
Best Local Similarity 100.0%; Pred. No. 0.017;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 480 GGVLAALAAAYCL 491

RESULT 328

ID AAR29868 standard; protein; 768 AA.

XX AAR29868;

XX 25-MAR-2003 (revised)

DT 26-APR-1993 (first entry)

XX HCV NS2-NS4 peptide N23N15A-1.

XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;

KW transcriptase; cDNA; primer; allele.

XX Hepatitis C virus.

XX EP518313-A2.

XX 16-DEC-1992.

XX 11-JUN-1992; 92EP-00109812.

XX 11-JUN-1991; 91JP-00139268.

PR 12-JUL-1991; 91JP-00172794.

PR 07-OCT-1991; 91JP-00287008.

PR 16-DEC-1991; 91JP-00332329.

PR 20-APR-1992; 92JP-00099957.

XX (MITU) MITSUBISHI KASEI CORP.

XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX WPI; 1992-417213/51.

DR N-PSDB; AAQ32499.

XX New hepatitis C virus gene and its encoded protein - used for diagnosing

PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 178-82; 305pp; English.

XX The sequences given in AAR29852-70 are encoded by various clones which
CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C
CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and
CC AAR29843-51). These RNA sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAQ32578-79. The sequences were then amplified
CC using primer pairs. The cDNA sequences isolated represent different
CC alleles of the same region of the HCV gene. Sequence comparisons of these
CC clones showed that it is possible for a patient to carry more than one
CC HCV strain at one time. See also AAQ32436. (Updated on 25-MAR-2003 to
CC correct PN field.)

XX Sequence 768 AA;

Query Match 10.2%; Score 12; DB 2; Length 768;

Best Local Similarity 100.0%; Pred. No. 0.017;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 513 GGVLAALAAAYCL 524

RESULT 329

AAR29850

ID AAR29850 standard; protein; 768 AA.

XX AAR29850;

XX 25-MAR-2003 (revised)

DT 26-APR-1993 (first entry)

XX HCV NS2-NS4 peptide N23N15.

XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;

KW transcriptase; cDNA; primer; allele.

XX Hepatitis C virus.

XX Key Location/Qualifiers

FT Misc-difference 19 /label= Pro, Leu

FT Misc-difference 24 /label= His, Arg

FT Misc-difference 60 /label= Thr, Ala

FT Misc-difference 71 /label= Ser, Thr

FT Misc-difference 88 /label= Lys, Arg

FT Misc-difference 140 /label= Ile, Val

FT Misc-difference 241 /label= Thr, Ile

FT Misc-difference 274 /label= Val, Ile

FT Misc-difference 381 /label= Pro, Ser

FT Misc-difference 475 /label= Tyr, His

FT Misc-difference 483 /label= Gln, Lys

FT Misc-difference 485 /label= Thr, Ala

FT Misc-difference 493 /label= Tyr, Phe

FT Misc-difference 496 /label= Thr, Ala

XX EP518313-A2.

XX 16-DEC-1992.

XX 11-JUN-1992; 92EP-00109812.

XX 11-JUN-1991; 91JP-00139268.

PR 12-JUL-1991; 91JP-00172794.

PR 07-OCT-1991; 91JP-00287008.

PR 16-DEC-1991; 91JP-00332329.

PR 20-APR-1992; 92JP-00099957.

XX (MITU) MITSUBISHI KASEI CORP.

XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX WPI; 1992-417213/51.

DR N-PSDB; AAQ32481.

XX New hepatitis C virus gene and its encoded protein - used for diagnosing

PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 134-38; 305pp; English.

XX The sequences given in AAR29860, AAR29559-60 and AAR29843-51 were encoded
CC by clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV)
CC gene of the invention. These sequences were isolated from the serum of a
CC patient suffering from hepatitis C (HC). The NS2-NS4 RNA sequences were
CC converted into cDNA using transcriptase in the presence of one of the
CC primer sequences given in AAR29553-64. The cDNA sequences were then
CC amplified using primer pairs. The cDNA sequences isolated represent
CC different alleles of the same region of the HCV gene. Sequence
CC comparisons of these clones showed that it is possible for a patient to
CC carry more than one HCV strain at one time. See also AAR29843. (Updated
CC on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 768 AA;

Query Match 10.2%; Score 12; DB 2; Length 768;
Best Local Similarity 100.0%; Pred. No. 0.017; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

QY 18 GGVLAALAAAYCL 29
|||||
Db 513 GGVLAALAAAYCL 524

RESULT 330
AAR29869
ID AAR29869 standard; protein; 768 AA.

XX AAR29869;
XX
XX 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX
XX HCV NS2-NS4 peptide N23N15B-1.
DE
XX
XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
KW transcriptase; cDNA; primer; allele.

XX Hepatitis C virus.

XX EP518313-A2.

XX 16-DEC-1992.

XX 11-JUN-1992; 92EP-00109812.

XX 11-JUN-1991; 91JP-00139268.

XX 12-JUL-1991; 91JP-00172794.

XX 07-OCT-1991; 91JP-00287008.

XX 16-DEC-1991; 91JP-00332329.

XX 20-APR-1992; 92JP-00099957.

XX (MITU) MITSUBISHI KASEI CORP.

XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX WPI; 1992-417213/51.

XX N-PSDB; AAR29559-60 and

XX New hepatitis C virus gene and its encoded protein - used for diagnosing

XX and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 182-86; 305pp; English.

XX The sequences given in AAR29852-70 are encoded by various clones which

XX were used in the isolation of the NS2-NS4 regions of the Hepatitis C

XX Virus (HCV) gene of the invention (see also AAR29860, AAR29559-60 and

XX AAR29843-51). These RNA sequences were isolated from the serum of a

XX patient suffering from hepatitis C (HC). The isolated RNA sequences were

CC clones showed that it is possible for a patient to carry more than one
CC HCV strain at one time. See also AAR29843. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 768 AA;

Query Match 10.2%; Score 12; DB 2; Length 768;
Best Local Similarity 100.0%; Pred. No. 0.017; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

QY 18 GGVLAALAAAYCL 29
|||||
Db 513 GGVLAALAAAYCL 524

RESULT 331
AAR82693
ID AAR82693 standard; protein; 916 AA.

XX AAR82693;

XX 16-OCT-2003 (revised)

DT 11-NOV-1996 (first entry)

XX HCV partial proteinase.

XX proteinase; hepatitis C virus; screening; inhibitor; proteolytic;

XX identification; cleavage.

XX Hepatitis C virus; Virus.

XX JP07184648-A.

XX 25-JUL-1995.

XX 05-FEB-1993; 93JP-00018854.

XX 07-FEB-1992; 92JP-00022657.

XX 18-SEP-1992; 92JP-00249240.

XX 04-DEC-1992; 92JP-00325303.

XX (KAEN/) KAENNO K.

XX (SUMQ) SUMITOMO METAL IND LTD.

XX (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.

XX WPI; 1995-287962/38.

XX N-PSDB; AAR03959.

XX An HCV proteinase active substance - which has activity as an anti-HCV

XX agent and can be used to screen for proteinase inhibitors.

XX Claim 3; Page 24-27; 52pp; Japanese.

XX The present sequence is that of a partial proteinase isolated from

XX Hepatitis C virus (HCV). The proteinase can be used as an anti-HCV agent.

XX It can also be used to screen cpds. for their ability to inhibit its

XX proteolytic activity. In this way proteinase inhibitors can be

XX identified. (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 916 AA;

Query Match 10.2%; Score 12; DB 2; Length 916;
Best Local Similarity 100.0%; Pred. No. 0.02; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;

QY 18 GGVLAALAAAYCL 29
|||||
Db 673 GGVLAALAAAYCL 684

RESULT 332
AAR82696
ID AAR82696 standard; protein; 923 AA.

FT	Misc-difference	178	/label= Met, Val
FT	Misc-difference	190	/label= Met, Ala
FT	Misc-difference	197	/label= Ala, Val
FT	Misc-difference	207	/label= Leu, Phe
FT	Misc-difference	208	/label= Met, Val
FT	Misc-difference	218	/label= Val, Ile
FT	Misc-difference	220	/label= Asp, Val
FT	Misc-difference	223	/label= Thr, Ala
FT	Misc-difference	227	/label= Asp, His
FT	Misc-difference	244	/label= Ala, Val
FT	Misc-difference	258	/label= Asp, Glu
FT	Misc-difference	267	/label= Leu, Ser
FT	Misc-difference	277	/label= Asn, Arg, Lys
FT	Misc-difference	279	/label= Ile, Leu
FT	Misc-difference	281	/label= Leu, Phe
FT	Misc-difference	343	/label= Ile, Val
FT	Misc-difference	385	/label= Gly, Arg
FT	Misc-difference	439	/label= Leu, Pro
FT	Misc-difference	444	/label= His, Arg
FT	Misc-difference	480	/label= Thr, Ala
FT	Misc-difference	491	/label= Ser, Thr
FT	Misc-difference	508	/label= Lys, Arg
FT	Misc-difference	560	/label= Ile, Val
FT	Misc-difference	661	/label= Thr, Ile
FT	Misc-difference	694	/label= Val, Ile
FT	Misc-difference	801	/label= Pro, Ser
FT	Misc-difference	895	/label= Tyr, His
FT	Misc-difference	903	/label= Glu, Lys
FT	Misc-difference	905	/label= Thr, Ala
FT	Misc-difference	913	/label= Tyr, Phe
FT	Misc-difference	916	/label= Thr, Ala
XX	EP5103113-A2.		
XX	16-DEC-1992.		
PP	11-JUN-1992;	92BP-00109812.	
XX	11-JUN-1991;	91JP-00139268.	
PR	12-JUL-1991;	91JP-00172794.	
PR	07-OCT-1991;	91JP-00287008.	
PR	16-DEC-1991;	91JP-00332329.	

```

PR 20-APR-1992; 92JP-00099957.
XX
XX (MITU) MITSUBISHI KASEI CORP.
XX
XX PI Seki M, Honda Y, Takahaashi K, Murakami T, Teranishi Y, Hayashi N;
XX
XX WPI; 1992-417213/51.
XX
XX DR N-PSDB; AAQ32442.
XX
XX PT New hepatitis C virus gene and its encoded protein - used for diagnosing
XX PT and vaccinating against hepatitis C virus infections.
XX
XX ES Disclosure; Page 139-45; 305pp; English.
XX
XX CC The sequences given in AAR29660, AAR29559-60 and AAR29843-51 were encoded
XX CC by clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV)
XX CC gene of the invention. These sequences were isolated from the serum of a
XX CC patient suffering from hepatitis C (HC). The NS2-NS4 RNA sequences were
XX CC converted into cDNA using transcriptase in the presence of one of the
XX CC primer sequences given in AAQ32553-64. The cDNA sequences were then
XX CC amplified using primer pairs. The cDNA sequences isolated represent
XX CC different alleles of the same region of the HCV gene. Sequence
XX CC comparisons of these clones showed that it is possible for a patient to
XX CC carry more than one HCV strain at one time. See also AAQ32436. (Updated
XX CC on 25-MAR-2003 to correct FN field.)
XX
XX SQ Sequence 1188 AA;
XX
XX Query Match 10.2%; Score 12; DB 2; Length 1188;
XX Best Local Similarity 100.0%; Pred. NO. 0.024;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCL 29
XX |||||
XX Db 933 GGVLAALAAAYCL 944
XX
XX RESULT 335
XX AAR29870
XX ID ID AAR29870 standard; protein; 1188 AA.
XX AC AAR29870;
XX
XX DT 25-MAR-2003 (revised)
XX DT 26-APR-1993 (first entry)
XX
XX DE HCV NS2-NS4 peptide MK25N15-1.
XX
XX KW Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
XX KW transcriptase; cDNA; primer; allele.
XX
XX OS Hepatitis C virus.
XX
XX EN EP518313-A2.
XX PD 16-DEC-1992.
XX
XX PF 11-JUN-1992; 92EP-00109812.
XX
XX PR 11-JUN-1991; 91JP-00139268.
XX PR 12-JUL-1991; 91JP-00172794.
XX PR 07-OCT-1991; 91JP-00287008.
XX PR 16-DEC-1991; 91JP-00332329.
XX PR 20-APR-1992; 92JP-00099957.
XX
XX PA (MITU) MITSUBISHI KASEI CORP.
XX
XX PI Seki M, Honda Y, Takahaashi K, Murakami T, Teranishi Y, Hayashi N;
XX
XX DR WPI; 1992-417213/51.
XX DR N-PSDB; AAQ32501.
XX
XX PT New hepatitis C virus gene and its encoded protein - used for diagnosing

```

PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 186-92; 305pp; English.

PS The sequences given in AAR29852-70 are encoded by various clones which

CC were used in the isolation of the NS2-NS4 regions of the Hepatitis C

CC Virus (HCV) gene of the invention (see also AAR29660, AAR29559-60 and

CC AAR2943-51). These RNA sequences were isolated from the serum of a

CC patient suffering from hepatitis C (HC). The isolated RNA sequences were

CC converted into cDNA using transcriptase in the presence of one of the

CC primer sequences given in AAR29578-79. The sequences were then amplified

CC using primer pairs. The cDNA sequences isolated represent different

CC alleles of the same region of the HCV gene. Sequence comparisons of these

CC clones showed that it is possible for a patient to carry more than one

CC HCV strain at one time. See also AAR32436. (Updated on 25-MAR-2003 to

CC correct PN field.)

XX Sequence 1188 AA;

SQ

Query Match 10.2%; Score 12; DB 2; Length 1188;

Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29

Db 933 GGVLAAALAAAYCL 944

RESULT 336

AAR12599

ID AAR12599 standard; protein; 1250 AA.

XX

AC AAR12599;

XX

XX 25-MAR-2003 (revised)

DT 17-SEP-1991 (first entry)

XX

XX Portion of PT-NANBH viral non-structural protein.

DE

XX post-transfusional non-A, non-B hepatitis; virus; vaccine; ss.

XX

XX Non-A.

OS non-B hepatitis virus.

OS

XX GB223245-A.

PN

XX 26-JUN-1991.

PD

XX 17-DEC-1990; 90GB-00027250.

XX

XX 18-DEC-1989; 89GB-00028562.

PR

XX 27-FEB-1990; 90GB-00004414.

PR

XX 03-MAR-1990; 90GB-00004814.

XX

XX (WELL) WELLCOME FOUND LTD.

PA

XX (HIGH/) HIGHFIELD P E.

XX

XX Highfield PE, Rodgers BC, Tedder RS, Barbara JAJ;

XX

XX WPI; 1991-187584/26.

DR

XX N-PSDB; AAR12241.

DR

XX Post-transfusional non-A non-B hepatitis poly:peptide(s) - and also DNA

PT

XX and antibodies used in diagnostic assays and in vaccines.

XX

XX Claim 1; Page 88-97; 108pp; English.

PS

XX The sequence was deduced from a non-structural (3') coding region

CC

CC sequence isolated from serum of humans infected by the PT-NANBH virus.

CC

CC The polypeptide is an antigenic portion of the virus and will be useful

CC

CC in the development of vaccines for inducing immunity in man to PT-NANBH.

CC

CC The invention covers PT-NANBH viral polypeptides having an amino acid

CC

CC sequence at least 90 per cent homologous with the sequence given here, or

CC

CC antigenic fragments of such homologous sequences. See also AAR12236-40

CC and AAR12242. (Updated on 25-MAR-2003 to correct PA field.)

XX

XX Sequence 1250 AA;

SQ

Query Match 10.2%; Score 12; DB 2; Length 1250;

Best Local Similarity 100.0%; Pred. No. 0.026;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29

Db 107 GGVLAAALAAAYCL 118

RESULT 337

AEA32849

ID AEA32849 standard; protein; 1394 AA.

XX

AC AEA32849;

XX

XX 28-JUL-2005 (first entry)

DT

XX

XX Modified NS3-5A genotype 1b protein, SEQ ID 3.

DE

XX NS5B; replicon; hepatitis C virus infection; antiinflammatory;

XX

XX hepatotropic; virucide.

KW

XX Synthetic.

OS

XX

XX Key Location/Qualifiers

PH

FT Misc-difference 904

FT /label= Val, Ala

FT

FT Misc-difference 1215

FT /label= Asn, Ser

FT

XX WO2005047463-A2.

PN

XX 26-MAY-2005.

PD

XX 03-NOV-2004; 2004WO-US036575.

XX

XX 05-NOV-2003; 2003US-0517605P.

XX

XX (MERI) MERCK & CO INC.

PA

XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.

XX

XX Ludmerer SW, Graham DJ, Lafemina RL, Flores OA, Pizzuti M;

PI

XX Traboni C;

PI

XX WPI; 2005-372359/38.

DR

XX N-PSDB; AEA32850.

DR

XX Enhancing ability of genotype 2b NS5B sequence to function in replicon

PT

XX for producing replicons containing functional genotype 2b NS5B, and

PT

XX measuring ability of compound to inhibit replicon activity, useful for

PT

XX treating hepatitis C.

XX

XX Claim 4; SEQ ID NO 3; 46pp; English.

PS

XX The invention relates to a novel method for enhancing the ability of a

CC

CC genotype 2b NS5B sequence to function in a replicon. The method comprises

CC

CC altering either or both the genotype 2b NS5B sequence to encode one or

CC

CC more adaptive mutations, or a genotype 1b NS4B sequence to encode an

CC

CC adaptive mutation of alanine at position 218 of a fully defined 261 amino

CC

CC acid (AEA32874) sequence given in the specification. The invention

CC

CC further comprises: a method for producing a chimeric replicon, comprising

CC

CC replacing substantially all of an NS5B sequence of a hepatitis C virus

CC

CC (HCV) replicon encoding a fully defined 1394 amino acid (AEA32849)

CC

CC sequence, with a genotype 2b NS5B encoding nucleic acid sequence; a

CC

CC chimeric replicon comprising an NS3-5A sequence of a genotype 1b replicon

CC

CC or a modified 2b NS3-5A sequence of a genotype 1b replicon, where NS4B

CC

CC contains a Val-218-Ala modification, and substantially all of a genotype

CC

CC 2b NS5B encoding nucleic acid sequence; and a recombinant cell comprising

CC a replicon of one of the methods or chimeric replicon, where the replicon
 CC is expressed in the cell. The method has virucide activity. The method is
 CC useful for enhancing the ability of a genotype 2b NS5B sequence to
 CC function in a replicon. The chimeric replicon and recombinant cell are
 CC useful for measuring the ability of a compound to inhibit replicon
 CC activity. The compounds tested can be used to treat or inhibit the onset
 CC of hepatitis C virus (HCV) infection in a patient. The method is useful
 CC for producing replicons containing functional genotype 2b NS5B. This
 CC sequence represents a NS3-5A genotype 1b protein used in the novel method
 CC of the invention.

XX
 XX Sequence 1394 AA;
 SQ
 Query Match 10.2%; Score 12; DB 9; Length 1394;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 639 GGVLAALAAAYCL 650

RESULT 338
 AAR29533
 ID AAR29533 standard; protein; 1411 AA.

XX
 AC AAR29533;
 AC
 XX 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)
 DT
 XX HCV NS4-NS5 peptide 1530U.

DE
 XX Clone; polypeptide; NS4-NS5; Hepatitis C; Virus; HCV; serum; HC;
 KW transcriptase; cDNA; primer; allele; core; region; upstream; hydrophilic;
 KW turn structure; alpha helix; beta sheet; antigen; determinant; antiserum.

XX Hepatitis C virus.

OS
 XX EP518313-A2.
 PN
 XX 16-DEC-1992.

PD
 XX 11-JUN-1992; 92EP-00109812.

PF
 XX 11-JUN-1991; 91JP-00139268.
 PR
 XX 12-JUL-1991; 91JP-00172794.
 PR
 XX 07-OCT-1991; 91JP-00287008.
 PR
 XX 16-DEC-1991; 91JP-00332329.
 PR
 XX 20-APR-1992; 92JP-00099957.

XX (MITU) MITSUBISHI KASEI CORP.

PA
 XX
 PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
 XX
 WPI; 1992-417213/51.
 DR
 DR N-PSDB; AAQ32443.

XX New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 216-24; 305pp; English.

XX The sequences given in AAR29871-81 and AAR29533 are encoded by various
 CC clones of the NS4-NS5 regions of the Hepatitis C Virus (HCV) gene of the
 CC invention. These NS4-NS5 RNA sequences were isolated from the serum of a
 CC patient suffering from hepatitis C (HC). The isolated RNA sequences were
 CC converted into cDNA using transcriptase in the presence of one of the
 CC primer sequences given in AAQ32565-77. The sequences were then amplified
 CC using primer pairs. The cDNA sequences isolated represent different
 CC alleles of the same region of the HCV gene. Sequence analysis shows that
 CC these clones represent the core region and some upstream sequences of
 CC HCV. These polypeptides are thought to contain a highly hydrophilic

CC region which can adopt a "turn structure" which is not an alpha helix or
 CC a beta sheet. These polypeptides are thought to act as antigen
 CC determinants and are highly reactive with antiserum raised against HCV-
 CC associated antigens. See also AAQ32436. (Updated on 25-MAR-2003 to
 CC correct PN field.)

XX Sequence 1411 AA;

SQ
 Query Match 10.2%; Score 12; DB 2; Length 1411;
 Best Local Similarity 100.0%; Pred. No. 0.028;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 65 GGVLAALAAAYCL 76

RESULT 339
 AAB36932
 ID AAB36932 standard; protein; 1736 AA.

XX
 AC AAB36932;

XX
 DT 06-AUG-2003 (revised)
 DT 27-FEB-2001 (first entry)

XX Hepatitis C virus tTA gene fused to polyprotein region.

XX NS3-5 polyprotein; virus activity; surrogate cell.

XX Hepatitis C virus.

OS Synthetic.

XX WO2000066623-A2.

XX 09-NOV-2000.

XX 01-MAY-2000; 2000WO-CA000496.

XX 04-MAY-1999; 99US-0132360P.

XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.

XX Pellerin C, Lamarre D;

XX WPI; 2001-007207/01.

DR N-PSDB; AAC83408.

XX Surrogate cell-based system for assaying hepatitis C virus (HCV) NS3
 PT protease activity, comprises chimeric DNA containing a transactivator
 PT domain fused downstream of a HCV DNA molecule encoding NS3-5 polyprotein.

XX Claim 9; Page 63-68; 70pp; English.

XX The present invention relates to a surrogate cell-based system to
 CC evaluate the activity of hepatitis C virus (HCV). The system involves a
 CC chimeric DNA molecule with a non-cytopathic expression system, a HCV
 CC recombinant DNA molecule encoding NS3-5 polyprotein linked to the
 CC expression system and a transactivator domain fused downstream of HCV DNA
 CC molecule and a second chimeric DNA molecule encoding a reporter gene. The
 CC surrogate cell-based system is useful for evaluating the activity of
 CC hepatitis C virus NS3 protease and is useful for identifying potential
 CC inhibitors of HCV NS3 protease activity. (Updated on 06-AUG-2003 to
 CC correct OS field.)

XX Sequence 1736 AA;

Query Match 10.2%; Score 12; DB 4; Length 1736;
 Best Local Similarity 100.0%; Pred. No. 0.034;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||

```
Db          641 GGVLAALAAAYCL 652
RESULT 340
AAO18001
ID AAO18001 standard; protein; 1985 AA.
AC AAO18001;
XX
XX 30-AUG-2002 (first entry)
DT
XX
XX Hepatitis C virus NS3 proteinase/helicase.
XX
XX Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;
KW virucide; hepatotropic; gene therapy; anti-viral; enzyme.
XX
XX Hepatitis C virus.
OS
XX
XX WO200238793-A2.
PN
XX
XX 16-MAY-2002.
PD
XX
XX 02-NOV-2001; 2001WO-US046350.
PF
XX
XX 07-NOV-2000; 2000US-0245866P.
PR
XX
XX (ANAD-) ANADYS PHARM INC.
PA
XX
XX Bichko V;
PI
XX
XX WPI; 2002-490082/52.
DR
XX
XX N-PSDB; AAL47276.
DR
XX
XX Novel nucleic acid encoding replication competent recombinant hepatitis C
PT virus genome useful for screening anti-hepatitis C virus therapeutics and
PT for vaccine development.
XX
XX Claim 6; Page 50-51; 85pp; English.
XX
XX The present invention provides protein and coding sequences from
CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and
CC able to replicate efficiently when transfected into a susceptible cell
CC line without reducing the growth rate of the cell line by more than 10
CC fold. The sequences are useful for screening for anti-HCV therapeutics,
CC for detecting antibodies to HCV in a biological sample such as blood,
CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,
CC for deriving authentic HCV components such as replication-component non-
CC infectious, replication-defective infection-component, and replication-
CC defective non-infectious HCV, in gene therapy or gene vaccination
CC targeted to hepatic tissue for treating an animal infected or susceptible
CC to HCV infection and for studying HCV infection and propagation. The
CC present sequence is the HCV NS3 proteinase/helicase
XX
XX Sequence 1985 AA;
SQ
Query Match 10.2%; Score 12; DB 5; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
DB 639 GGVLAALAAAYCL 650
RESULT 341
AAE15729
ID AAE15729 standard; protein; 1985 AA.
XX
XX AAE15729;
XX
XX 12-MAR-2002 (first entry)
DT
XX
XX Hepatitis C virus (HCV) replBartMan polyprotein variant A1174S.
DE
```

```
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW replBartMan; mutant; mutain; variant.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 1174 /note= "Wild type Ala substituted with Ser"
FT
XX
XX WO200189364-A2.
PN
XX
XX 29-NOV-2001.
PD
XX
XX 23-MAY-2001; 2001WO-US016822.
PF
XX
XX 23-MAY-2000; 2000US-00576989.
PR
XX
XX (UNIW ) UNIV WASHINGTON.
PA
XX
XX Rice CM, Blight KJ;
PI
XX
XX WPI; 2002-066755/09.
DR
XX
XX Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
XX Claim 14; Page; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polyprotein coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting for
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) replBartMan polyprotein
CC variant. Note: The present sequence is not shown in the specification but
CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
CC specification
XX
SQ Sequence 1985 AA;
Query Match 10.2%; Score 12; DB 5; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
DB 639 GGVLAALAAAYCL 650
```

RESULT 342
AAE15731
ID AAE15731 standard; protein; 1985 AA.
XX AC AAE15731;
XX DT 12-MAR-2002 (first entry)
XX DE Hepatitis C virus (HCV) repibBartMan polypeptide variant S1172P.
XX KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
XX KW repibBartMan; mutant; mucin; variant.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 1172 /note= "Wild type Ser substituted with Pro"
XX WO200189364-A2.
XX PD 29-NOV-2001.
XX PF 23-MAY-2001; 2001WO-US016822.
XX PR 23-MAY-2000; 2000US-00576989.
XX PA (UNIV) UNIV WASHINGTON.
XX PI Rice CM, Blight KJ;
XX DR WPI; 2002-066755/09.
XX PT Hepatitis C virus variants having greater transfection efficiency and
XX PT ability to survive subpassage, useful as a vaccine for immunizing primate
XX PT to the virus, comprise non-naturally occurring viral sequences.
XX PS Claim 14; Page; 174pp; English.
XX CC The invention relates to Hepatitis C virus (HCV) variants which include
XX CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX CC variants that have a transfection efficiency and ability to survive
XX CC subpassage greater than HCV that have wild-type polypeptide coding
XX CC regions. The polynucleotides of the invention are useful for identifying
XX CC a cell line that is permissive for infection with HCV and detecting
XX CC replication of HCV in cells of the cell line. They are also useful for
XX CC testing a compound for anti-viral properties and for inhibiting HCV
XX CC infection. They are also useful for the generation of defined HCV virus
XX CC stocks to develop in vitro and in vivo assays for virus neutralisation,
XX CC attachment, penetration and entry, structure/function studies on HCV
XX CC proteins and RNA elements and identification of new antiviral targets, a
XX CC systematic survey of cell culture systems and conditions to identify
XX CC those that support wild-type and variant HCV RNA replication and particle
XX CC release, production of adaptive HCV variants capable of more efficiency
XX CC replication in cell culture, production of HCV variants with altered
XX CC tissue or species tropism, establishment of alternative animal models for
XX CC inhibitor evaluation including those supporting HCV variant replication,
XX CC development of cell-free HCV replication assays, production of
XX CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
XX CC derivatives as possible vaccine candidates, engineering of attenuated or
XX CC defective HCV derivatives for expression of heterologous gene products
XX CC for gene therapy and vaccine applications and for utilisation of the HCV
XX CC glycoproteins for targeted delivery of therapeutic agents to the liver
XX CC or other cell types with appropriate receptors. Vaccine comprising these
XX CC sequences is useful for inducing immunoprotection to HCV in a primate.
XX CC The present sequence is Hepatitis C virus (HCV) repibBartMan polypeptide
XX CC variant. Note: The present sequence is not shown in the specification but
XX CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
XX CC specification

SQ Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 5; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 639 GGVLAALAAAYCL 650
|||||

RESULT 343

AAE15720
ID AAE15720 standard; protein; 1985 AA.

XX AC AAE15720;

XX DT 12-MAR-2002 (first entry)

XX DE Hepatitis C virus (HCV) adaptive replicon VI protein.

XX KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
XX KW adaptive replicon VI.

XX OS Hepatitis C virus.

XX PN WO200189364-A2.

XX PD 29-NOV-2001.

XX PF 23-MAY-2001; 2001WO-US016822.

XX PR 23-MAY-2000; 2000US-00576989.

XX PA (UNIV) UNIV WASHINGTON.

XX PI Rice CM, Blight KJ;

XX DR WPI; 2002-066755/09.

XX DR N-PSDB; AAD25324.

XX PT Hepatitis C virus variants having greater transfection efficiency and
XX PT ability to survive subpassage, useful as a vaccine for immunizing primate
XX PT to the virus, comprise non-naturally occurring viral sequences.

XX PS Example 1; Page 87-88; 174pp; English.

XX CC The invention relates to Hepatitis C virus (HCV) variants which include
XX CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX CC variants that have a transfection efficiency and ability to survive
XX CC subpassage greater than HCV that have wild-type polypeptide coding
XX CC regions. The polynucleotides of the invention are useful for identifying
XX CC a cell line that is permissive for infection with HCV and detecting
XX CC replication of HCV in cells of the cell line. They are also useful for
XX CC testing a compound for anti-viral properties and for inhibiting HCV
XX CC infection. They are also useful for the generation of defined HCV virus
XX CC stocks to develop in vitro and in vivo assays for virus neutralisation,
XX CC attachment, penetration and entry, structure/function studies on HCV
XX CC proteins and RNA elements and identification of new antiviral targets, a
XX CC systematic survey of cell culture systems and conditions to identify
XX CC those that support wild-type and variant HCV RNA replication and particle
XX CC release, production of adaptive HCV variants capable of more efficiency
XX CC replication in cell culture, production of HCV variants with altered
XX CC tissue or species tropism, establishment of alternative animal models for
XX CC inhibitor evaluation including those supporting HCV variant replication,
XX CC development of cell-free HCV replication assays, production of
XX CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
XX CC derivatives as possible vaccine candidates, engineering of attenuated or
XX CC defective HCV derivatives for expression of heterologous gene products
XX CC for gene therapy and vaccine applications and for utilisation of the HCV
XX CC glycoproteins for targeted delivery of therapeutic agents to the liver
XX CC or other cell types with appropriate receptors. Vaccine comprising these
XX CC sequences is useful for inducing immunoprotection to HCV in a primate.
XX CC The present sequence is Hepatitis C virus (HCV) repibBartMan polypeptide
XX CC variant. Note: The present sequence is not shown in the specification but
XX CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
XX CC specification

CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) adaptive replicon VI
CC protein
XX
SQ Sequence 1985 AA;
Query Match 10.2%; Score 12; DB 5; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 639 GGVLAALAAAYCL 650
RESULT 344
AAE15717
ID AAE15717 standard; protein; 1985 AA.
XX
AC AAE15717;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) replBartMan polyprotein.
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW replBartMan.
XX
XX Hepatitis C virus.
OS
PN WO200189364-A2.
XX
PD 29-NOV-2001.
XX
PP 23-MAY-2001; 2001WO-US016822.
XX
PR 23-MAY-2000; 2000US-00576989.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Rice CM, Blight KJ;
XX
XX WPI; 2002-066755/09.
DR N-ESDB; AAD25322.
XX
PT Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
PS Claim 14; Page 65; 174pp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or

CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is Hepatitis C virus (HCV) replBartMan polyprotein
XX
SQ Sequence 1985 AA;
Query Match 10.2%; Score 12; DB 5; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 639 GGVLAALAAAYCL 650
RESULT 345
AAE15727
ID AAE15727 standard; protein; 1985 AA.
XX
AC AAE15727;
XX
DT 12-MAR-2002 (first entry)
XX
DE Hepatitis C virus (HCV) replBartMan polyprotein variant S11791.
XX
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
KW replBartMan; mutant; mutein; variant.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 1179 /note= "Wild type Ser substituted with Ile"
FT
XX
XX WO200189364-A2.
PN
XX 29-NOV-2001.
PD
XX 23-MAY-2001; 2001WO-US016822.
PP
XX 23-MAY-2000; 2000US-00576989.
PR
XX (UNIW) UNIV WASHINGTON.
PA
XX Rice CM, Blight KJ;
PI
XX WPI; 2002-066755/09.
DR
XX Hepatitis C virus variants having greater transfection efficiency and
XX ability to survive subpassage, useful as a vaccine for immunizing primate
XX to the virus, comprise non-naturally occurring viral sequences.
PS Claim 14; Page; 174pp; English.
XX
CC The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle

CC release, production of adaptive HCV variants capable of more efficiency
 CC replication in cell culture, production of HCV variants with altered
 CC tissue or species tropism, establishment of alternative animal models for
 CC inhibitor evaluation including those supporting HCV variant replication,
 CC development of cell-free HCV replication assays, production of
 CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
 CC derivatives as possible vaccine candidates, engineering of attenuated or
 CC defective HCV derivatives for expression of heterologous gene products
 CC for gene therapy and vaccine applications and for utilisation of the HCV
 CC glycoproteins for targeted delivery of therapeutic agents to the liver
 CC or other cell types with appropriate receptors. Vaccine comprising these
 CC sequences is useful for inducing immunoprotection to HCV in a primate.
 CC The present sequence is Hepatitis C virus (HCV) replBartMan polypeptide
 CC variant. Note: the present sequence is not shown in the specification but
 CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
 CC specification
 CC
 CC SQ Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 5; Length 1985;
 Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 639 GGVLAALAAAYCL 650

RESULT 346

AAE15728
 ID AAE15728 standard; protein; 1985 AA.

XX AC AAE15728;

XX DT 12-MAR-2002 (first entry)

XX DE Hepatitis C virus (HCV) replBartMan polypeptide variant R1164G.

XX KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
 XX gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
 XX replBartMan; mutant; mutein; variant.

XX OS Hepatitis C virus.
 XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1164 /note= "Wild type Arg substituted with Gly"

XX FN WO200189364-A2.

XX PD 29-NOV-2001.

XX PF 23-MAY-2001; 2001WO-US016822.

XX PR 23-MAY-2000; 2000US-00576989.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Rice CM, Blight KJ;

XX DR WPI; 2002-066755/09.

XX PT Hepatitis C virus variants having greater transfection efficiency and
 XX ability to survive subpassage, useful as a vaccine for immunizing primate
 XX to the virus, comprise non-naturally occurring viral sequences.

XX PS Claim 14; Page; 174pp; English.

XX CC The invention relates to Hepatitis C virus (HCV) variants which include
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
 CC variants that have a transfection efficiency and ability to survive
 CC subpassage greater than HCV that have wild-type polypeptide coding

CC regions. The polynucleotides of the invention are useful for identifying
 CC a cell line that is permissive for infection with HCV and detecting
 CC replication of HCV in cells of the cell line. They are also useful for
 CC testing a compound for anti-viral properties and for inhibiting HCV
 CC infection. They are also useful for the generation of defined HCV virus
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,
 CC attachment, penetration and entry, structure/function studies on HCV
 CC proteins and RNA elements and identification of new antiviral targets, a
 CC systematic survey of cell culture systems and conditions to identify
 CC those that support wild-type and variant HCV RNA replication and particle
 CC release, production of adaptive HCV variants capable of more efficiency
 CC replication in cell culture, production of HCV variants with altered
 CC tissue or species tropism, establishment of alternative animal models for
 CC inhibitor evaluation including those supporting HCV variant replication,
 CC development of cell-free HCV replication assays, production of
 CC immunogenic HCV particles for vaccination, engineering of attenuated or
 CC defective HCV derivatives for expression of heterologous gene products
 CC for gene therapy and vaccine applications and for utilisation of the HCV
 CC glycoproteins for targeted delivery of therapeutic agents to the liver
 CC or other cell types with appropriate receptors. Vaccine comprising these
 CC sequences is useful for inducing immunoprotection to HCV in a primate.
 CC The present sequence is Hepatitis C virus (HCV) replBartMan polypeptide
 CC variant. Note: the present sequence is not shown in the specification but
 CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
 CC specification
 CC
 CC SQ Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 5; Length 1985;
 Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 639 GGVLAALAAAYCL 650

RESULT 347

AAE15722
 ID AAE15722 standard; protein; 1985 AA.

XX AC AAE15722;

XX DT 12-MAR-2002 (first entry)

XX DE Hepatitis C virus (HCV) adaptive replicon II mutant polypeptide.

XX KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
 XX gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
 XX adaptive replicon II; mutant; mutein; variant.

XX OS Hepatitis C virus.
 XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1164 /note= "Wild-type Arg substituted with Gly"

XX FN WO200189364-A2.

XX PD 29-NOV-2001.

XX PF 23-MAY-2001; 2001WO-US016822.

XX PR 23-MAY-2000; 2000US-00576989.

XX PA (UNIW) UNIV WASHINGTON.

XX PI Rice CM, Blight KJ;

XX DR WPI; 2002-066755/09.

XX N-PSDB; AAD25325.

XX Hepatitis C virus variants having greater transfection efficiency and
PT ability to survive subpassage, useful as a vaccine for immunizing primate
PT to the virus, comprise non-naturally occurring viral sequences.
XX
XX Example 1; Page 88-89; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV
CC variants that have a transfection efficiency and ability to survive
CC subpassage greater than HCV that have wild-type polypeptide coding
CC regions. The polynucleotides of the invention are useful for identifying
CC a cell line that is permissive for infection with HCV and detecting
CC replication of HCV in cells of the cell line. They are also useful for
CC testing a compound for anti-viral properties and for inhibiting HCV
CC infection. They are also useful for the generation of defined HCV virus
CC stocks to develop in vitro and in vivo assays for virus neutralisation,
CC attachment, penetration and entry, structure/function studies on HCV
CC proteins and RNA elements and identification of new antiviral targets, a
CC systematic survey of cell culture systems and conditions to identify
CC those that support wild-type and variant HCV RNA replication and particle
CC release, production of adaptive HCV variants capable of more efficiency
CC replication in cell culture, production of HCV variants with altered
CC tissue or species tropism, establishment of alternative animal models for
CC inhibitor evaluation including those supporting HCV variant replication,
CC development of cell-free HCV replication assays, production of
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV
CC derivatives as possible vaccine candidates, engineering of attenuated or
CC defective HCV derivatives for expression of heterologous gene products
CC for gene therapy and vaccine applications and for utilisation of the HCV
CC glycoproteins for targeted delivery of therapeutic agents to the liver
CC or other cell types with appropriate receptors. Vaccine comprising these
CC sequences is useful for inducing immunoprotection to HCV in a primate.
CC The present sequence is a Hepatitis C virus (HCV) adaptive replicon II
CC mutant protein. This sequence is generated by the mutation Arg to Gly at
CC position 1164 of HCVreplBartMan protein
XX
XX Sequence 1985 AA;
XX
XX Query Match 10.2%; Score 12; DB 5; Length 1985;
XX Best Local Similarity 100.0%; Pred. No. 0.038;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCL 29
XX |||||
XX Db 639 GGVLAALAAAYCL 650
XX
XX RESULT 348
XX AAEL5730
XX ID AAEL5730 standard; protein; 1985 AA.
XX AC AAEL5730;
XX
XX 12-MAR-2002 (first entry)
XX
XX Hepatitis C virus (HCV) replBartMan polypeptide variant S1172C.
XX
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;
XX gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;
XX replBartMan; mutant; mutein; variant.
XX
XX Hepatitis C virus.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 1172 /note= "Wild type Ser substituted with Cys"
XX
XX WO200189364-A2.
XX
XX 29-NOV-2001.
XX

PP 23-MAY-2001; 2001WO-US016822.
XX
XX 23-MAY-2000; 2000US-00576989.
XX
XX (UNIW) UNIV WASHINGTON.
XX
XX Rice CM, Blight KJ;
XX
XX WPT; 2002-066755/09.
XX
XX Hepatitis C virus variants having greater transfection efficiency and
XX ability to survive subpassage, useful as a vaccine for immunizing primate
XX to the virus, comprise non-naturally occurring viral sequences.
XX
XX Claim 14; Page; 174pp; English.
XX
XX The invention relates to Hepatitis C virus (HCV) variants which include
XX polynucleotides comprising non-naturally occurring HCV sequence and HCV
XX variants that have a transfection efficiency and ability to survive
XX subpassage greater than HCV that have wild-type polypeptide coding
XX regions. The polynucleotides of the invention are useful for identifying
XX a cell line that is permissive for infection with HCV and detecting
XX replication of HCV in cells of the cell line. They are also useful for
XX testing a compound for anti-viral properties and for inhibiting HCV
XX infection. They are also useful for the generation of defined HCV virus
XX stocks to develop in vitro and in vivo assays for virus neutralisation,
XX attachment, penetration and entry, structure/function studies on HCV
XX proteins and RNA elements and identification of new antiviral targets, a
XX systematic survey of cell culture systems and conditions to identify
XX those that support wild-type and variant HCV RNA replication and particle
XX release, production of adaptive HCV variants capable of more efficiency
XX replication in cell culture, production of HCV variants with altered
XX tissue or species tropism, establishment of alternative animal models for
XX inhibitor evaluation including those supporting HCV variant replication,
XX development of cell-free HCV replication assays, production of
XX immunogenic HCV particles for vaccination, engineering of attenuated HCV
XX derivatives as possible vaccine candidates, engineering of attenuated or
XX defective HCV derivatives for expression of heterologous gene products
XX for gene therapy and vaccine applications and for utilisation of the HCV
XX glycoproteins for targeted delivery of therapeutic agents to the liver
XX or other cell types with appropriate receptors. Vaccine comprising these
XX sequences is useful for inducing immunoprotection to HCV in a primate.
XX The present sequence is a Hepatitis C virus (HCV) replBartMan polypeptide
XX variant. Note: The present sequence is not shown in the specification but
XX is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the
XX specification
XX
XX Sequence 1985 AA;
XX
XX Query Match 10.2%; Score 12; DB 5; Length 1985;
XX Best Local Similarity 100.0%; Pred. No. 0.038;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 18 GGVLAALAAAYCL 29
XX |||||
XX Db 639 GGVLAALAAAYCL 650
XX
XX RESULT 349
XX ABU09574
XX ID ABU09574 standard; protein; 1985 AA.
XX AC ABU09574;
XX
XX 09-JUL-2003 (first entry)
XX
XX HCV Met-NS3-NS4A-NS5A-NS5B (inactivated).
XX
XX HCV; enzyme; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B;
XX adenoviral vector; HCV infection; vaccine; mutant; gene therapy;
XX protease; mutein.
XX
XX Hepatitis C virus.
XX

OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 1711..1713
 FT /note= "Wild-type Gly-Asp substituted by Ala-Ala-Gly"
 XX
 XX WO2003031588-A2.
 XX
 XX PD 17-APR-2003.
 XX
 XX PF 10-OCT-2002; 2002WO-US032512.
 XX
 XX PR 11-OCT-2001; 2001US-0328655P.
 XX
 XX PR 13-MAR-2002; 2002US-0363774P.
 XX
 XX PA (MERI) MERCK & CO INC.
 XX (RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.
 XX
 XX PI Emni EA, Kaslow DC, Bett AJ, Shiver JW, Nicosia A, Lahm A;
 XX PI Luzzago A, Cortese R, Colloca S;
 XX
 XX DR WPI; 2003-381708/36.
 XX DR N-PSDB; ACA61111, ACA61112, ACA61113, ACA61118, ACA61119.
 XX
 XX PT New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide,
 XX PT useful as a component of an adenovector or DNA plasmid vaccine for
 XX PT preventing or treating hepatitis C virus.
 XX
 XX PS Claim 1; Fig 1; 231pp; English.
 XX
 XX CC The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A
 XX CC -NS5B (NS stands for non-structural protein from HCV) polypeptide
 XX CC appearing as AB009574. The encoded polypeptide has sufficient protease
 XX CC activity to process itself to produce an NS5B protein that is
 XX CC enzymatically inactive. Also included are a cultured recombinant cell
 XX CC comprising the novel nucleic acid, making an adenovector (comprising: (a)
 XX CC producing an adenovirus genome plasmid comprising a gene expression
 XX CC cassette by homologous recombination between the novel nucleic acid and a
 XX CC nucleic acid comprising a first adenovirus region from base pair 1-450
 XX CC corresponding to either Ad5 or Ad6, a second adenovirus region from base
 XX CC pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541
 XX CC corresponding to Ad6, joined to the first region, a third adenovirus
 XX CC region from base pair 5549-28133 corresponding to Ad5 or from base pair
 XX CC 5542-28156 corresponding to Ad6, joined to the second region, a fourth
 XX CC adenovirus region from base pair 30818-33966 corresponding to Ad5 or from
 XX CC base pair 30789-33784 corresponding to Ad6, joined to the third region,
 XX CC and a fifth adenovirus region from base pair 33967-35935 corresponding to
 XX CC Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the
 XX CC fourth region; and (b) rescuing the recombinant adenovirus from the
 XX CC recombinant adenovirus plasmid), an adenoviral vector that is produced by
 XX CC method above appearing as ACA61113 which has a humanised version of the
 XX CC polynucleotide of the invention and encodes the HCV inactivated
 XX CC polypeptide, a recombinant nucleic acid comprising one or more Ad6
 XX CC regions and a region not present in Ad6, where at least one Ad6 region is
 XX CC selected from E1A, E1B, E2B, E2A, E4, L1, L2, L4 and L5. The nucleic acid
 XX CC is useful as a component of an adenoviral vector or a DNA plasmid vaccine
 XX CC providing a broad range of antigens for generating an HCV-specific cell
 XX CC mediated immune response. The nucleic acid may also be used in treating
 XX CC patients infected with HCV. The present sequence is the HCV polypeptide
 XX CC with an inactive NS5B protein

Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 6; Length 1985;
 Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 639 GGVLAALAAAYCL 650

RESULT 350

ABU09575
 ID ABU09575 standard; protein; 1985 AA.
 XX
 AC ABU09575;
 XX
 XX 09-JUL-2003 (first entry)
 XX
 XX DE HCV Met-NS3-NS4A-NS4B-NS5A-NS5B (active).
 XX
 XX KW HCV; enzyme; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B;
 XX KW adenoviral vector; HCV infection; vaccine; gene therapy; protease.
 XX
 XX OS Hepatitis C virus.
 XX
 XX PN WO2003031588-A2.
 XX
 XX PD 17-APR-2003.
 XX
 XX PF 10-OCT-2002; 2002WO-US032512.
 XX
 XX PR 11-OCT-2001; 2001US-0328655P.
 XX PR 13-MAR-2002; 2002US-0363774P.
 XX
 XX PA (MERI) MERCK & CO INC.
 XX (RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.
 XX
 XX PI Emni EA, Kaslow DC, Bett AJ, Shiver JW, Nicosia A, Lahm A;
 XX PI Luzzago A, Cortese R, Colloca S;
 XX
 XX XX WPI; 2003-381708/36.
 XX DR N-PSDB; ACA61114.
 XX
 XX PT New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide,
 XX PT useful as a component of an adenovector or DNA plasmid vaccine for
 XX PT preventing or treating hepatitis C virus.
 XX
 XX PS Example 1; Fig 5; 231pp; English.
 XX
 XX CC The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A
 XX CC -NS5B (NS stands for non-structural protein from HCV) polypeptide
 XX CC appearing as ABU09574. The encoded polypeptide has sufficient protease
 XX CC activity to process itself to produce an NS5B protein that is
 XX CC enzymatically inactive. Also included are a cultured recombinant cell
 XX CC comprising the novel nucleic acid, making an adenovector (comprising: (a)
 XX CC producing an adenovirus genome plasmid comprising a gene expression
 XX CC cassette by homologous recombination between the novel nucleic acid and a
 XX CC nucleic acid comprising a first adenovirus region from base pair 1-450
 XX CC corresponding to either Ad5 or Ad6, a second adenovirus region from base
 XX CC pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541
 XX CC corresponding to Ad6, joined to the first region, a third adenovirus
 XX CC region from base pair 5549-28133 corresponding to Ad5 or from base pair
 XX CC 5542-28156 corresponding to Ad6, joined to the second region, a fourth
 XX CC adenovirus region from base pair 30818-33966 corresponding to Ad5 or from
 XX CC base pair 30789-33784 corresponding to Ad6, joined to the third region,
 XX CC and a fifth adenovirus region from base pair 33967-35935 corresponding to
 XX CC Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the
 XX CC fourth region; and (b) rescuing the recombinant adenovirus from the
 XX CC recombinant adenovirus plasmid), an adenoviral vector that is produced by
 XX CC method above appearing as ACA61113 which has a humanised version of the
 XX CC polynucleotide of the invention and encodes the HCV inactivated
 XX CC polypeptide, a recombinant nucleic acid comprising one or more Ad6
 XX CC regions and a region not present in Ad6, where at least one Ad6 region is
 XX CC selected from E1A, E1B, E2B, E2A, E4, L1, L2, L4 and L5. The nucleic acid
 XX CC is useful as a component of an adenoviral vector or a DNA plasmid vaccine
 XX CC providing a broad range of antigens for generating an HCV-specific cell
 XX CC mediated immune response. The nucleic acid may also be used in treating
 XX CC patients infected with HCV. The present sequence is the wild-type HCV
 XX CC polypeptide with an active NS5B protein

Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 6; Length 1985;
 Best Local Similarity 100.0%; Pred. No. 0.038;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 639 GGVLAALAAAYCL 650

RESULT 351
ADJ57846
ID ADJ57846 standard; protein; 1985 AA.
AC ADJ57846;
XX
XX
DT 06-MAY-2004 (first entry)
XX
XX HCV replicon.
XX hepatitis C virus; HCV; Antiinflammatory; Hepatotropic; Virucide;
XX HCV replicon.
XX Unidentified.
XX OS
XX PN WO2004015131-A2.
XX XX
XX PD 19-FEB-2004.
XX XX
XX PF 12-AUG-2003; 2003WO-US025260.
XX XX
XX PR 12-AUG-2002; 2002US-0402661P.
XX XX
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX XX
XX PI Gao M, Lemm JA, O'boyle DR, Nower P, Rigat K, Sun J;
XX XX
XX WPI; 2004-180685/17.
XX N-PSDB; ADJ57845.
XX
XX Use of hepatitis C virus assays or reporter assays, e.g. identifying a
XX compound that inhibits hepatitis C virus RNA replication or identifying a
XX compound that modulates the activity of a gene of interest.
XX
XX Claim 4; SEQ ID NO 2; 45pp; English.
XX
XX The present invention relates to the use of hepatitis C virus (HCV)
XX assays for identifying a compound that inhibits HCV RNA replication and
XX reporter assays for identifying a compound that modulates the activity of
XX a gene of interest. The assays are useful for identifying a compound that
XX inhibits HCV RNA replication or for identifying a compound that modulates
XX the activity of a gene of interest. The HCV assay is useful for high
XX throughput screening that quantifies both the amount of HCV RNA
XX replication inhibitory activity associated with a test compound and the
XX amount of cytotoxicity associated with the test compound. The compound is
XX useful for treating hepatitis C infection. Assays of the invention have
XX distinct advantages when compared to qRT-PCR or other methods in that
XX assays of the invention may take place in situ in a detergent based crude
XX cell lysate, which requires no further preparation prior to performing
XX the assays. The assays do not also involve numerous manipulations to add
XX or subtract reagents after addition of test compounds and are desirably
XX based on a viral protein which is required by the HCV replicon for
XX replication. The present sequence represents a HCV replicon used in the
XX assay of the invention.
XX
XX Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 8; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 639 GGVLAALAAAYCL 650

RESULT 352
ADR38450
ID ADR38450 standard; protein; 1985 AA.
XX
XX AC ADR38450;
XX
XX DT 18-NOV-2004 (first entry)
XX
XX Hepatitis C virus (isolate BK) NS3-NS5B polyprotein mutant Seq 1.
XX DB
XX Hepatitis C virus; HCV; NS3; NS5A; non-structural protein;
XX virus replication activity; mutant; mutein.
XX KW
XX Hepatitis C virus (isolate BK).
XX OS Synthetic.
XX XX
XX Key Location/Qualifiers
XX Protein 2..632
XX /note= "NS3"
XX Misc-difference 471
XX /note= "Wild type residue substituted for Met,
XX corresponds to posn 470 of NS3"
XX FT
XX Protein 633..686
XX /note= "NS4A"
XX Protein 687..947
XX /note= "NS4B"
XX Protein 948..1394
XX /note= "NS5A"
XX Misc-difference 1179
XX /note= "Wild type residue substituted for Ile,
XX corresponds to posn 232 of NS5A"
XX FT
XX Protein 1395..1985
XX /note= "NS5B"
XX
XX WO2004074507-A2.
XX
XX 02-SEP-2004.
XX
XX 09-FEB-2004; 2004WO-US003726.
XX
XX 13-FEB-2003; 2003US-0447318P.
XX
XX (MERI) MERCK & CO INC.
XX
XX Grobler J, Flores O, Markel EJ;
XX
XX WPI; 2004-635590/61.
XX N-PSDB; ADR38452.
XX
XX Making Hepatitis C virus (HCV) replicon having increased replication
XX activity, useful in HCV research, comprises modifying HCV replicon
XX construct to encode an amino acid substitution at a position
XX corresponding to amino acid 470 of NS3.
XX
XX Claim 9; SEQ ID NO 1; 54pp; English.
XX
XX This invention relates to a novel method for producing a Hepatitis C
XX virus (HCV) replicon having an increased replication activity.
XX Specifically, it refers to modifying an HCV replicon construct to encode
XX an amino acid substitution in NS3 (a non-structural protein that along
XX with NS4A, NS4B, NS5A and NS5B make up the virus replication machinery
XX released in the form of a polyprotein). The present invention describes
XX an amino acid substitution at a position corresponding to amino acid 470
XX of NS3 alone, or in combination with, an isoleucine in a position
XX corresponding to amino acid 232 of NS5A that confers improved cell
XX culture replication activity compared to wild type HCV. The method is
XX useful for facilitating the identification of broadly efficacious
XX compounds against different HCV isolates and facilitating HCV research,
XX where compounds that inhibit HCV replication have research and
XX therapeutic applications in identifying overall efficacy and lack of
XX unacceptable toxicity. Accordingly, they can be used to treat or inhibit
XX the onset of HCV in a patient. This polypeptide sequence is the HCV NS3-
XX NS5B polyprotein based on HCV-BK with a 471M mutation (corresponds to

CC 470M of the NS3 protein) and 11791 mutation, given in an exemplification
CC of the invention.

XX Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 8; Length 1985;

XX Best Local Similarity 100.0%; Pred. No. 0.038; 0; Indels 0; Gaps 0;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

|||||

Db 639 GGVLAALAAAYCL 650

RESULT 353

ABE17075

ID ABE17075 standard; protein; 1985 AA.

XX ABE17075;

XX 22-SEP-2005 (first entry)

XX Hepatitis C virus isolate Con1 polyprotein, SEQ ID NO: 42 #1.

XX RNA virus; microorganism; hepatitis C virus infection; antiinflammatory;
KW hepatotropic; virucide; gastrointestinal disease; infection; polyprotein;
KW ds; gene.

XX Hepatitis C virus; isolate Con1.

XX Key Location/Qualifiers

FT Region 1..632

FT /note= "NS3 protein"

FT Region 633..686

FT /note= "NS4A protein"

FT Region 687..947

FT /note= "NS4B protein"

FT Region 948..1394

FT /note= "NS5A protein"

FT Region 1395..1985

FT /note= "NS5B protein"

XX US2005153281-A1.

XX 14-JUL-2005.

XX 06-DEC-2004; 2004US-00006313.

XX 23-DEC-1999; 99US-0171909P.

XX 23-DEC-2000; 2000US-00747419.

XX 27-SEP-2001; 2001US-0325236P.

XX 13-NOV-2001; 2001US-0338123P.

XX 27-SEP-2002; 2002US-00259275.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Lemon SM, Yi M;

XX WPI; 2005-496820/50.

XX N-PSDB; ABE17074.

XX Detecting replication competent hepatitis C virus (HCV) RNA by incubating
FT a vertebrate cell comprising an HCV RNA and detecting the detectable
PT marker, where its presence indicates the cell comprises a replication
PT competent HCV RNA.

XX Disclosure; SEQ ID NO 42; 142pp; English.

XX The present invention provides a replication competent hepatitis C virus
CC (HCV) RNA that includes a heterologous polynucleotide. The invention also
CC includes methods for modifying a HCV polynucleotide, selecting a
CC replication competent HCV polynucleotide, detecting a replication
CC competent HCV polynucleotide and identifying a compound that inhibits

CC replication of a HCV polynucleotide. The present sequence is hepatitis C
CC virus (HCV) isolate Con1 polyprotein. This sequence includes non-
CC structural replicative proteins which includes NS3 (a serine
CC proteinase/NTase/RNA helicase), NS4A (serine proteinase accessory
CC factor), NS4B, NS5A and NS5B (RNA-dependent RNA polymerase) and 3' non-
CC translated region (3' NTR). Note: The present sequence is the SEQ ID NO:
CC 42 which is given in the sequence listing. This sequence differs from the
CC SEQ ID NO: 42 shown in the disclosure (figure 24e) of the specification.

XX Sequence 1985 AA;

Query Match 10.2%; Score 12; DB 9; Length 1985;

XX Best Local Similarity 100.0%; Pred. No. 0.038;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

|||||

Db 639 GGVLAALAAAYCL 650

RESULT 354

ADD67963

ID ADD67963 standard; protein; 2063 AA.

XX ADD67963;

XX 15-JAN-2004 (first entry)

XX Hepatitis C virus polyprotein.

XX antiviral; hepatitis C virus; HCV; viral replication inhibitor;
KW replication competent HCV; 3' non-translated region; polyprotein.

XX Hepatitis C virus.

XX US2003125541-A1.

XX 03-JUL-2003.

XX 27-SEP-2002; 2002US-00259275.

XX 23-DEC-1999; 99US-0171909P.

XX 23-DEC-2000; 2000US-00747419.

XX 27-SEP-2001; 2001US-0325236P.

XX 13-NOV-2001; 2001US-0338123P.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Lemon SM, Yi M;

XX WPI; 2003-811006/76.

XX Identifying a compound that inhibits replication of a hepatitis C virus
PT (HCV) RNA comprises contacting a cell comprising a replication competent
PT HCV RNA containing a heterologous polynucleotide encoding a
PT transactivator, with a compound.

XX Disclosure; Fig 24; 95pp; English.

XX The invention describes a method of identifying a compound that inhibits
CC replication of a hepatitis C virus (HCV) RNA. The method comprises
CC contacting a cell comprising a replication competent HCV RNA containing a
CC heterologous polynucleotide having a first coding sequence encoding a
CC transactivator, with a compound. The method is useful for identifying a
CC compound that inhibits replication of HCV RNA. The kit is useful for
CC detecting replication competent HCV RNA. This is the amino acid sequence
CC of hepatitis C virus polyprotein.

XX Sequence 2063 AA;

Query Match 10.2%; Score 12; DB 7; Length 2063;

XX Best Local Similarity 100.0%; Pred. No. 0.039;

XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 639 GGVLAALAAAYCL 650

RESULT 355
AEB17116
ID AEB17116 standard; protein; 2063 AA.
XX
AC AEB17116;
XX
DT 22-SEP-2005 (first entry)
XX
DE Hepatitis C virus Isolate Con1 polyprotein, SEQ ID NO: 42 #2.
XX
KW RNA virus; microorganism; hepatitis C virus infection; antiinflammatory;
KW hepatotropic; virucide; gastrointestinal disease; infection; polyprotein;
KW ds; gene.
XX
OS Hepatitis C virus; Isolate Con1.

Key Location/Qualifiers
FH 1..632
FT Region /note= "NS3 protein"
FT Region 633..686
FT Region /note= "NS4A protein"
FT Region 687..947
FT Region /note= "NS4B protein"
FT Region 948..1394
FT Region /note= "NS5A protein"
FT Region 1395..1985
FT Region /note= "NS5B protein"
FT Region 1986..2063
FT Region /note= "3' NTR protein"
FT Misc-difference 1986 /note= "Encoded by TGA"
FT Misc-difference 1990 /note= "Encoded by TAA"
FT Misc-difference 2041 /note= "Encoded by TGA"
FT Misc-difference 2043 /note= "Encoded by TGA"
FT Misc-difference 2055 /note= "Encoded by TGA"

US2005153281-A1.
14-JUL-2005.
06-DEC-2004; 2004US-00006313.
23-DEC-1999; 99US-0171909P.
23-DEC-2000; 2000US-00747419.
27-SEP-2001; 2001US-0325236P.
13-NOV-2001; 2001US-0338123P.
27-SEP-2002; 2002US-00259275.
(TEXA) UNIV TEXAS SYSTEM.
Lemon SM, Yi M;
WPI; 2005-496820/50.
N-PSDB; AEB17074.

Detecting replication competent hepatitis C virus (HCV) RNA by incubating a vertebrate cell comprising an HCV RNA and detecting the detectable marker, where its presence indicates the cell comprises a replication competent HCV RNA.

Disclosure; SEQ ID NO 42; 142pp; English.

The present invention provides a replication competent hepatitis C virus

CC (HCV) RNA that includes a heterologous polynucleotide. The invention also includes methods for modifying a HCV polynucleotide, selecting a replication competent HCV polynucleotide, detecting a replication competent HCV polynucleotide and identifying a compound that inhibits replication of a HCV polynucleotide. The present sequence is hepatitis C virus (HCV) Isolate Con1 polyprotein. This sequence includes non-structural replicative proteins which includes NS3 (a serine proteinase/NTase/RNA helicase), NS4A (serine proteinase accessory factor), NS4B, NS5A and NS5B (RNA-dependent RNA polymerase) and 3' non-translated region (3' NTR). Note: The present sequence is the SEQ ID NO: 42 which is shown in the disclosure (figure 24e) of the specification. This sequence differs from the SEQ ID NO: 42 given in the sequence listing of the specification.

XX
SQ Sequence 2063 AA;
Query Match 10.2%; Score 12; DB 9; Length 2063;
Best Local Similarity 100.0%; Pred. No. 0.039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 639 GGVLAALAAAYCL 650

RESULT 356
AAW01680
ID AAW01680 standard; protein; 2201 AA.
XX
AC AAW01680;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 19-AUG-1997 (first entry)
XX
DE HCV NS2-NS5B non-structural protein.
XX
KW Hepatitis C virus; HCV; non-structural protein; NS5B; polyprotein;
KW RNA-dependent RNA polymerase; RdRp; terminal nucleotidyl transferase;
KW Tase; method; assay; in vitro activity; therapy; inhibitor.
XX
OS Hepatitis C virus; strain BK.

WO9637619-A1.
28-NOV-1996.
24-MAY-1996; 96WO-IT000106.
25-MAY-1995; 95IT-RM000343.
(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
De Francesco R, Tomei L, Behrens S;
WPI; 1997-021225/02.

Reproducing enzymatic activities of HCV in vitro - using sequences contg. NS5B for RNA-dependent RNA polymerase and terminal nucleotidyl transferase activities.

Example 1; Page 26-34; 49pp; English.

A novel method for reproducing in vitro the RNA-dependent RNA polymerase (RdRp) activity or the terminal nucleotidyl transferase (TNTase) activity encoded by hepatitis C virus (HCV), is characterised in that sequences contg. NS5B are used in the mixture. The method is used for assaying in vitro the activities of RdRp and TNTase encoded by HCV in order to identify, for therapeutic purposes, compounds that inhibit these enzymatic activities and therefore might interfere with the replication of the HCV. The present sequence comprises amino acids 810-3010 of the HCV polyprotein and corresponds to NS2-NS5B proteins. cDNA encoding this protein was cloned between the NcoI and HindIII sites of pBlueBacII to

CC form pBac25. Another expression plasmid, pBac5B (containing cDNA encoding
CC amino acids 2420-3010 of HCV; see AAW01679) was also constructed.
CC Extracts of Bac25- or Bac5B-infected Sf9 cells contain a novel magnesium-
CC dependent enzymatic activity that catalyses de novo RNA synthesis. This
CC activity was shown to be dependent on the presence of RNA, but
CC independent of an added primer or of the origin of the input RNA
CC molecule. As the products generated by extracts of Sf9 cells infected
CC with either Bac25 or Bac5 appeared to be identical, the experiments
CC indicated that the observed RdRp activity is encoded by the HCV NS5B
CC protein. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 17-OCT
CC -2003 to standardise OS field)
XX
XX

SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 2; Length 2201;

Best Local Similarity 100.0%; Pred. No. 0.041;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

DB 855 GGVLAALAAAYCL 866

RESULT 357

ABG30601

ID ABG30601 standard; protein; 2201 AA.

XX

AC ABG30601;

XX 21-OCT-2002 (first entry)

XX Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #10.

XX Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; muten.
XX

OS Hepatitis C virus.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 882

FT /label= Arg, Lys

FT Misc-difference 2183

FT /note= "Wild type Met substituted by Thr"

XX WO200252015-A2.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-CA001843.

XX 22-DEC-2000; 2000US-0257857P.

XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.

XX Kukolj G, Pause A;

XX WPI; 2002-575382/61.

XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.

XX Claim 3; Page; 140pp; English.

XX The invention describes a self-replicating hepatitis C virus (HCV)
CC polynucleotide molecule comprising a 5'-non translated region (NTR),
CC where guanine at position 1 is substituted for adenine, a HCV polypeptide
CC region coding for a HCV polypeptide; and a 3'-NTR region. The self-
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at

CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence represents a mutant of the hepatitis
CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
CC This sequence does not appear in the specification but has been created
CC from the wild type sequence shown in ABG30580 using information given in
CC the claims of the invention
XX

SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;

Best Local Similarity 100.0%; Pred. No. 0.041;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

DB 855 GGVLAALAAAYCL 866

RESULT 358

ABG30584

ID ABG30584 standard; protein; 2201 AA.

XX

AC ABG30584;

XX 21-OCT-2002 (first entry)

XX Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #4.

XX Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; muten.
XX

OS Hepatitis C virus.

XX

XX Key Location/Qualifiers

FT Misc-difference 393

FT /note= "Wild type Glu substituted by Gly"

FT Misc-difference 1175

FT /note= "Wild type Ile substituted by Val"

FT Misc-difference 1233

FT /note= "Wild type Gly substituted by Cys"

FT Misc-difference 2183

FT /note= "Wild type Met substituted by Thr"

XX WO200252015-A2.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-CA001843.

XX 22-DEC-2000; 2000US-0257857P.

XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.

XX Kukolj G, Pause A;

XX WPI; 2002-575382/61.

XX N-PSDB; ABK88576.

XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.

XX Claim 3; Page 85-95; 140pp; English.

XX The invention describes a self-replicating hepatitis C virus (HCV)
CC polynucleotide molecule comprising a 5'-non translated region (NTR),
CC where guanine at position 1 is substituted for adenine, a HCV polypeptide
CC region coding for a HCV polypeptide; and a 3'-NTR region. The self-
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at

CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at
CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence is encoded by the hepatitis C virus
CC replicon APGK12 and contains the viral protease NS2/3, protease complex
CC NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note: this
CC sequence has been created from replicon APGK12 shown in ABG30581
XX
XX
SQ Sequence 2201 AA;
Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866
RESULT 359
ABG30591
ID ABG30591 standard; protein; 2201 AA.
XX AC
XX ABG30591;
DT 21-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #3.
XX
XX Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutetin.
XX
XX Hepatitis C virus.
OS Synthetic.
OS
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 751 /note= "Wild type Ser substituted by Gly"
FT Misc-difference 882 /label= Arg, Lys
FT
XX WO200252015-A2.
XX
XX
XX 04-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-CA001843.
XX
XX 22-DEC-2000; 2000US-0257857P.
XX
XX (BOBH) BOEHRINGER INGELHEIM CANADA LTD.
XX
XX Kukolj G, Pause A;
PI
XX WPI; 2002-575382/61.
DR
XX
XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.
XX
XX Claim 3; Page; 140pp; English.
PS
XX
XX The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polypeptide
XX region coding for a HCV polypeptide; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at

CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence represents a mutant of the hepatitis
CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
CC This sequence does not appear in the specification but has been created
CC from the wild type sequence shown in ABG30580 using information given in
CC the claims of the invention
XX
XX
SQ Sequence 2201 AA;
Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866
RESULT 360
ABG30600
ID ABG30600 standard; protein; 2201 AA.
XX AC
XX ABG30600;
DT 21-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #9.
XX
XX Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutetin.
XX
XX Hepatitis C virus.
OS Synthetic.
OS
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 882 /label= Arg, Lys
FT Misc-difference 1357 /note= "Wild type Pro substituted by Leu"
FT
XX WO200252015-A2.
XX
XX
XX 04-JUL-2002.
XX
XX 20-DEC-2001; 2001WO-CA001843.
XX
XX 22-DEC-2000; 2000US-0257857P.
XX
XX (BOBH) BOEHRINGER INGELHEIM CANADA LTD.
XX
XX Kukolj G, Pause A;
PI
XX WPI; 2002-575382/61.
DR
XX
XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.
XX
XX Claim 3; Page; 140pp; English.
PS
XX
XX The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polypeptide
XX region coding for a HCV polypeptide; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to

CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence represents a mutant of the hepatitis
CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
CC This sequence does not appear in the specification but has been created
CC from the wild type sequence shown in ABG30580 using information given in
CC the claims of the invention
XX

SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
855 GGVLAALAAAYCL 866

RESULT 361

ABG30581
ID ABG30581 standard; protein; 2201 AA.

XX AC ABG30581;

XX DT 21-OCT-2002 (first entry)

XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #1.

XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B.

XX OS Hepatitis C virus.

XX PN WO200252015-A2.

XX PD 04-JUL-2002.

XX PF 20-DEC-2001; 2001WO-CA001843.

XX PR 22-DEC-2000; 2000US-0257857P.

XX PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.

XX PI Kukolj G, Pause A;

XX DR WPI; 2002-575382/61.

XX DR N-PSDB; ABK88573.

XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.

XX PS Disclosure; Page 49-58; 140pp; English.

XX The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polypeptide
XX region coding for a HCV polypeptide; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to
XX existing systems comprising a self-replicating HCV RNA molecule that, in
XX conjunction with mutations in the HCV non-structural region, such as the
XX G(2042)C/R mutations, transduces and/or replicates with greater
XX efficiency. This amino acid sequence is encoded by the hepatitis C virus
XX replicon APGK12 and contains the viral protease NS2/3, the protease complex
XX NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B

SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
855 GGVLAALAAAYCL 866

RESULT 362

ABG30586
ID ABG30586 standard; protein; 2201 AA.

XX AC ABG30586;

XX DT 21-OCT-2002 (first entry)

XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #6.

XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutein.

XX OS Hepatitis C virus.

XX FH Key Location/Qualifiers

XX FT Misc-difference 326 /note= "Wild type Arg substituted by Lys"

XX FT Misc-difference 751 /note= "Wild type Ser substituted by Gly"

XX FT Misc-difference 882 /note= "Wild type Lys substituted by Arg"

XX FT Misc-difference 1184 /note= "Wild type Thr substituted by Ala"

XX FT Misc-difference 1233 /note= "Wild type Gly substituted by Cys"

XX FT Misc-difference 1346 /note= "Wild type Leu substituted by Pro"

XX FT Misc-difference 1357 /note= "Wild type Pro substituted by Leu"

XX PN WO200252015-A2.

XX PD 04-JUL-2002.

XX PF 20-DEC-2001; 2001WO-CA001843.

XX PR 22-DEC-2000; 2000US-0257857P.

XX PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.

XX PI Kukolj G, Pause A;

XX DR WPI; 2002-575382/61.

XX DR N-PSDB; ABK88578.

XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.

XX PS Disclosure; Page 106-116; 140pp; English.

XX The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polypeptide
XX region coding for a HCV polypeptide; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to
XX existing systems comprising a self-replicating HCV RNA molecule that, in

CC conjunction with mutations in the HCV non-structural region, such as the
CC G(2042)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence is encoded by the hepatitis C virus
CC replicon APGK12 and contains the viral protease NS2/3, protease complex
CC NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note: this
CC sequence has been created from replicon APGK12 shown in ABG30581
XX
XX
SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 855 GGVLAALAAAYCL 866

RESULT 363

ABG30593

ID ABG30593 standard; protein; 2201 AA.

XX AC ABG30593;

XX DT 21-OCT-2002 (first entry)

XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #4.

XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; muten.
XX OS Hepatitis C virus.
XX OS Synthetic.

XX FH Key Location/Qualifiers
XX FT Misc-difference 882
XX FT Misc-difference 892 /label= Arg, Lys
XX FT Misc-difference 892 /note= "Wild type Leu substituted by Phe"
XX FN WO200252015-A2.
XX PD 04-JUL-2002.
XX PF 20-DEC-2001; 2001WO-CA001843.
XX PR 22-DEC-2000; 2000US-0257857P.
XX PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX PI Kukolj G, Pause A;
XX DR WPI; 2002-575382/61.
XX DR N-PSDB; ABK88574.

XX PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.
XX PS Claim 3; Page; 140pp; English.

XX CC The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polypeptide
XX region coding for a HCV polypeptide; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to
XX existing systems comprising a self-replicating HCV RNA molecule that, in
XX conjunction with mutations in the HCV non-structural region, such as the
XX G(2042)C/R mutations, transduces and/or replicates with greater
XX efficiency. This amino acid sequence represents a mutant of the hepatitis

CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
CC This sequence does not appear in the specification but has been created
CC from the wild type sequence shown in ABG30580 using information given in
CC the claims of the invention
XX
XX
SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 855 GGVLAALAAAYCL 866

RESULT 364

ABG30582

ID ABG30582 standard; protein; 2201 AA.

XX AC ABG30582;

XX DT 21-OCT-2002 (first entry)

XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #2.

XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; muten.
XX OS Hepatitis C virus.
XX OS Synthetic.

XX FH Key Location/Qualifiers
XX FT Misc-difference 882
XX FT Misc-difference 1233 /note= "Wild type Lys substituted by Lys or Arg"
XX FT Misc-difference 892 /note= "Wild type Gly substituted by Cys"
XX FN WO200252015-A2.
XX PD 04-JUL-2002.
XX PF 20-DEC-2001; 2001WO-CA001843.
XX PR 22-DEC-2000; 2000US-0257857P.
XX PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX PI Kukolj G, Pause A;
XX DR WPI; 2002-575382/61.
XX DR N-PSDB; ABK88574.

XX PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.
XX PS Disclosure; Page 59-69; 140pp; English.

XX CC The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polypeptide
XX region coding for a HCV polypeptide; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to
XX existing systems comprising a self-replicating HCV RNA molecule that, in
XX conjunction with mutations in the HCV non-structural region, such as the
XX G(2042)C/R mutations, transduces and/or replicates with greater
XX efficiency. This amino acid sequence is encoded by the hepatitis C virus
XX replicon APGK12 and contains the viral protease NS2/3, protease complex

CC NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note: this
CC sequence has been created from replicon APGK12 shown in ABG30581
XX
SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866
|||||

RESULT 365
ABG30580
ID ABG30580 standard; protein; 2201 AA.
XX AC ABG30580;
XX
DT 21-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #9.
XX
KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW cell culture replication; NS2/3; NS3/4; NS3; NS5B.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 882 /note= "Encoded by ARG"
FT
FN WO200252015-A2.
XX
PD 04-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-CA001843.
XX
PR 22-DEC-2000; 2000US-0257857P.
XX
PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX
PI Kukolj G, Pause A;
XX
DR WPI; 2002-575382/61.
XX
PS New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.
XX
XX Disclosure; Page 69-74; 140pp; English.

The invention describes a self-replicating hepatitis C virus (HCV)
polynucleotide molecule comprising a 5'-non translated region (NTR),
where guanine at position 1 is substituted for adenine, a HCV polyprotein
region coding for a HCV polyprotein; and a 3'-NTR region. The self-
replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
potential inhibitors of HCV replication. The HCV RNA molecule is also
useful for efficiently establishing cell culture replication. The self-
replicating polynucleotide molecule contains a 5'-NTR, where G at
position 1 is substituted for A, and therefore provides an alternative to
existing systems comprising a self-replicating HCV RNA molecule that, in
conjunction with mutations in the HCV non-structural region, such as the
G(2042)C/R mutations, transduces and/or replicates with greater
efficiency. This amino acid sequence represents a mutant of the hepatitis
C virus replicon APGK12 and contains the viral protease NS2/3, protease
complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
this sequence does not appear in the specification but has been created
from the wild type sequence shown in ABG30580 using information given in
the claims of the invention

XX
SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866
|||||

RESULT 366
ABG30602
ID ABG30602 standard; protein; 2201 AA.
XX AC ABG30602;
XX
DT 21-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #11.
XX
KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutetin.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT Misc-difference 393 /note= "Wild type Glu substituted by Gly"
FT Misc-difference 882 /label= Arg, Lys
FT
FN WO200252015-A2.
XX
PD 04-JUL-2002.
XX
PF 20-DEC-2001; 2001WO-CA001843.
XX
PR 22-DEC-2000; 2000US-0257857P.
XX
PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX
PI Kukolj G, Pause A;
XX
DR WPI; 2002-575382/61.
XX
PS New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.
XX
XX Claim 4; Page; 140pp; English.

The invention describes a self-replicating hepatitis C virus (HCV)
polynucleotide molecule comprising a 5'-non translated region (NTR),
where guanine at position 1 is substituted for adenine, a HCV polyprotein
region coding for a HCV polyprotein; and a 3'-NTR region. The self-
replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
potential inhibitors of HCV replication. The HCV RNA molecule is also
useful for efficiently establishing cell culture replication. The self-
replicating polynucleotide molecule contains a 5'-NTR, where G at
position 1 is substituted for A, and therefore provides an alternative to
existing systems comprising a self-replicating HCV RNA molecule that, in
conjunction with mutations in the HCV non-structural region, such as the
G(2042)C/R mutations, transduces and/or replicates with greater
efficiency. This amino acid sequence represents a mutant of the hepatitis
C virus replicon APGK12 and contains the viral protease NS2/3, protease
complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
this sequence does not appear in the specification but has been created
from the wild type sequence shown in ABG30580 using information given in
the claims of the invention

XX
SQ Sequence 2201 AA;

Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      18 GGVLAALAAAYCL 29
Db      855 GGVLAALAAAYCL 866

RESULT 367
ABG30587
ID      ABG30587 standard; protein; 2201 AA.
XX
AC      ABG30587;
XX
DT      21-OCT-2002 (first entry)
XX
DE      Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #7.
XX
KW      Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW      cell culture replication; NS2/3; NS3/4; NS3; NS5B.
XX
OS      Hepatitis C virus.
XX
XX
XX      Key      Location/Qualifiers
FT      Misc-difference 326
FT      /note= "Wild type Arg substituted by Lys"
FT      Misc-difference 882
FT      /label= Arg, Lys
XX
PN      WO200252015-A2.
XX
XX      04-JUL-2002.
XX
XX      20-DEC-2001; 2001WO-CA001843.
XX
XX      22-DEC-2000; 2000US-0257857P.
XX
XX      (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX
XX      Kukolj G, Pause A;
XX
XX      WPI; 2002-575382/61.
XX
XX      New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX      possess enhanced transduction or replication efficiency, useful for
XX      evaluating potential inhibitors of HCV replication.
XX
XX      Discloure; Page 120-129; 140pp; English.
XX
XX      The invention describes a self-replicating hepatitis C virus (HCV)
XX      polynucleotide molecule comprising a 5'-non translated region (NTR),
XX      where guanine at position 1 is substituted for adenine, a HCV polyprotein
XX      region coding for a HCV polyprotein; and a 3'-NTR region. The self-
XX      replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX      potential inhibitors of HCV replication. The HCV RNA molecule is also
XX      useful for efficiently establishing cell culture replication. The self-
XX      replicating systems comprising a self-replicating HCV RNA molecule that, in
XX      conjunction with mutations in the HCV non-structural region, such as the
XX      G(2042)C/R mutations, transduces and/or replicates with greater
XX      efficiency. This amino acid sequence represents a mutant of the hepatitis
XX      C virus replicon APGK12 and contains the viral protease NS2/3, protease
XX      complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
XX      This sequence does not appear in the specification but has been created
XX      from the wild type sequence shown in ABG30580 using information given in
XX      the claims of the invention
XX
XX      Sequence 2201 AA;

Query Match      10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      855 GGVLAALAAAYCL 866

RESULT 368
ABG30589
ID      ABG30589 standard; protein; 2201 AA.
XX
AC      ABG30589;
XX
DT      21-OCT-2002 (first entry)
XX
DE      Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #1.
XX
KW      Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
KW      cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutein.
XX
OS      Hepatitis C virus.
XX
XX
XX      Key      Location/Qualifiers
FT      Misc-difference 326
FT      /note= "Wild type Arg substituted by Lys"
FT      Misc-difference 882
FT      /label= Arg, Lys
XX
PN      WO200252015-A2.
XX
XX      04-JUL-2002.
XX
XX      20-DEC-2001; 2001WO-CA001843.
XX
XX      22-DEC-2000; 2000US-0257857P.
XX
XX      (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX
XX      Kukolj G, Pause A;
XX
XX      WPI; 2002-575382/61.
XX
XX      New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX      possess enhanced transduction or replication efficiency, useful for
XX      evaluating potential inhibitors of HCV replication.
XX
XX      Discloure; Page 120-129; 140pp; English.
XX
XX      The invention describes a self-replicating hepatitis C virus (HCV)
XX      polynucleotide molecule comprising a 5'-non translated region (NTR),
XX      where guanine at position 1 is substituted for adenine, a HCV polyprotein
XX      region coding for a HCV polyprotein; and a 3'-NTR region. The self-
XX      replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX      potential inhibitors of HCV replication. The HCV RNA molecule is also
XX      useful for efficiently establishing cell culture replication. The self-
XX      replicating systems comprising a self-replicating HCV RNA molecule that, in
XX      conjunction with mutations in the HCV non-structural region, such as the
XX      G(2042)C/R mutations, transduces and/or replicates with greater
XX      efficiency. This amino acid sequence is encoded by the hepatitis C virus
XX      replicon APGK12 and contains the viral protease NS2/3, protease complex
XX      NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B
XX
XX      Sequence 2201 AA;

Query Match      10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      855 GGVLAALAAAYCL 866

RESULT 369
ABG30599
ID      ABG30599 standard; protein; 2201 AA.
XX
AC      ABG30599;
```

XX DT 21-OCT-2002 (first entry)
XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #8.
XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutain.
XX OS Hepatitis C virus.
XX FH Synthetic.
XX FH Key Location/Qualifiers
FT Misc-difference 882 /label= Arg, Lys
FT Misc-difference 1346 /note= "Wild type Leu substituted by Pro"
XX FT
XX FT
XX PN WO200252015-A2.
XX PD 04-JUL-2002.
XX PF 20-DEC-2001; 2001WO-CA001843.
XX PR 22-DEC-2000; 2000US-0257857P.
XX PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX PI Kukolj G, Pause A;
XX WPI; 2002-575382/61.
XX DR
XX PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.
XX PS Claim 3; Page; 140pp; English.
XX CC The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polyprotein
XX region coding for a HCV polyprotein; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to
XX existing systems comprising a self-replicating HCV RNA molecule that, in
XX conjunction with mutations in the HCV non-structural region, such as the
XX G(2042)C/R mutations, transduces and/or replicates with greater
XX efficiency. This amino acid sequence represents a mutant of the hepatitis
XX C virus replicon APGK12 and contains the viral protease NS2/3, protease
XX complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
XX This sequence does not appear in the specification but has been created
XX from the wild type sequence shown in ABG30580 using information given in
XX the claims of the invention
XX SQ Sequence 2201 AA;
Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866
RESULT 370
ABG30585
ID ABG30585 standard; protein; 2201 AA.
XX AC
XX ABG30585;
XX

DT 21-OCT-2002 (first entry)
XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #5.
XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutain.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
FT Misc-difference 339 /note= "Wild type Ser substituted by Gly"
FT Misc-difference 393 /note= "Wild type Glu substituted by Gly"
FT Misc-difference 892 /note= "Wild type Leu substituted by Phe"
FT Misc-difference 1233 /note= "Wild type Gly substituted by Arg"
FT Misc-difference 1595 /note= "Wild type Ser substituted by Pro"
XX PN WO200252015-A2.
XX PD 04-JUL-2002.
XX PF 20-DEC-2001; 2001WO-CA001843.
XX PR 22-DEC-2000; 2000US-0257857P.
XX PA (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX PI Kukolj G, Pause A;
XX WPI; 2002-575382/61.
XX DR N-PSDB; ABK88577.
XX PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX possess enhanced transduction or replication efficiency, useful for
XX evaluating potential inhibitors of HCV replication.
XX PS Disclosure; Page 96-106; 140pp; English.
XX CC The invention describes a self-replicating hepatitis C virus (HCV)
XX polynucleotide molecule comprising a 5'-non translated region (NTR),
XX where guanine at position 1 is substituted for adenine, a HCV polyprotein
XX region coding for a HCV polyprotein; and a 3'-NTR region. The self-
XX replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX potential inhibitors of HCV replication. The HCV RNA molecule is also
XX useful for efficiently establishing cell culture replication. The self-
XX replicating polynucleotide molecule contains a 5'-NTR, where G at
XX position 1 is substituted for A, and therefore provides an alternative to
XX existing systems comprising a self-replicating HCV RNA molecule that, in
XX conjunction with mutations in the HCV non-structural region, such as the
XX G(2042)C/R mutations, transduces and/or replicates with greater
XX efficiency. This amino acid sequence is encoded by the hepatitis C virus
XX replicon APGK12 and contains the viral protease NS2/3, protease complex
XX NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note: this
XX sequence has been created from replicon APGK12 shown in ABG30581
XX SQ Sequence 2201 AA;
Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866
RESULT 371
ABG30594
ID ABG30594 standard; protein; 2201 AA.

```
XX AC ABG30594;
XX XX
XX DT 21-OCT-2002 (first entry)
XX XX
XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #5.
XX XX
XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX KW cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutein.
XX XX
XX OS Hepatitis C virus.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 882
XX FT /label= Arg, Lys
XX FT
XX FT Misc-difference 1175
XX FT /note= "Wild type Ile substituted by Val"
XX XX
XX PN WO200252015-A2.
XX XX
XX PD 04-JUL-2002.
XX XX
XX PF 20-DEC-2001; 2001WO-CA001843.
XX XX
XX PR 22-DEC-2000; 2000US-0257857P.
XX XX
XX PA (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX XX
XX PI Kukolj G, Pause A;
XX XX
XX DR WPI; 2002-575382/61.
XX XX
XX PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX PT possess enhanced transduction or replication efficiency, useful for
XX PT evaluating potential inhibitors of HCV replication.
XX XX
XX PS Claim 3; Page; 140pp; English.
XX XX
XX CC The invention describes a self-replicating hepatitis C virus (HCV)
XX CC polynucleotide molecule comprising a 5'-non translated region (NTR),
XX CC where guanine at position 1 is substituted for adenine, a HCV polyprotein
XX CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-
XX CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX CC potential inhibitors of HCV replication. The HCV RNA molecule is also
XX CC useful for efficiently establishing cell culture replication. The self-
XX CC replicating polynucleotide molecule contains a 5'-NTR, where G at
XX CC position 1 is substituted for A, and therefore provides an alternative to
XX CC existing systems comprising a self-replicating HCV RNA molecule that, in
XX CC conjunction with mutations in the HCV non-structural region, such as the
XX CC G(2042)C/R mutations, transduces and/or replicates with greater
XX CC efficiency. This amino acid sequence represents a mutant of the hepatitis
XX CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
XX CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
XX CC This sequence does not appear in the specification but has been created
XX CC from the wild type sequence shown in ABG30580 using information given in
XX CC the claims of the invention
XX XX
XX SQ Sequence 2201 AA;
XX XX
XX Query Match 10.2%; Score 12; DB 5; Length 2201;
XX Best Local Similarity 100.0%; Pred. No. 0.041;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCL 29
XX |||||
XX Db 855 GGVLAALAAAYCL 866
XX
XX RESULT 372
XX ABG30598
XX ID ABG30598 standard; protein; 2201 AA.
XX XX
```

```
AC ABG30598;
XX XX
XX DT 21-OCT-2002 (first entry)
XX XX
XX DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #7.
XX XX
XX KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
XX KW cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutein.
XX XX
XX OS Hepatitis C virus.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 882
XX FT /label= Arg, Lys
XX FT
XX FT Misc-difference 1595
XX FT /note= "Wild type Ser substituted by Pro"
XX XX
XX PN WO200252015-A2.
XX XX
XX PD 04-JUL-2002.
XX XX
XX PF 20-DEC-2001; 2001WO-CA001843.
XX XX
XX PR 22-DEC-2000; 2000US-0257857P.
XX XX
XX PA (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.
XX XX
XX PI Kukolj G, Pause A;
XX XX
XX DR WPI; 2002-575382/61.
XX XX
XX PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which
XX PT possess enhanced transduction or replication efficiency, useful for
XX PT evaluating potential inhibitors of HCV replication.
XX XX
XX PS Claim 3; Page; 140pp; English.
XX XX
XX CC The invention describes a self-replicating hepatitis C virus (HCV)
XX CC polynucleotide molecule comprising a 5'-non translated region (NTR),
XX CC where guanine at position 1 is substituted for adenine, a HCV polyprotein
XX CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-
XX CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
XX CC potential inhibitors of HCV replication. The HCV RNA molecule is also
XX CC useful for efficiently establishing cell culture replication. The self-
XX CC replicating polynucleotide molecule contains a 5'-NTR, where G at
XX CC position 1 is substituted for A, and therefore provides an alternative to
XX CC existing systems comprising a self-replicating HCV RNA molecule that, in
XX CC conjunction with mutations in the HCV non-structural region, such as the
XX CC G(2042)C/R mutations, transduces and/or replicates with greater
XX CC efficiency. This amino acid sequence represents a mutant of the hepatitis
XX CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
XX CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
XX CC This sequence does not appear in the specification but has been created
XX CC from the wild type sequence shown in ABG30580 using information given in
XX CC the claims of the invention
XX XX
XX SQ Sequence 2201 AA;
XX XX
XX Query Match 10.2%; Score 12; DB 5; Length 2201;
XX Best Local Similarity 100.0%; Pred. No. 0.041;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCL 29
XX |||||
XX Db 855 GGVLAALAAAYCL 866
XX
XX RESULT 373
XX ABG30595
XX ID ABG30595 standard; protein; 2201 AA.
XX XX
XX AC ABG30595;
```

XX 21-OCT-2002 (first entry)
DT Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #6.
XX Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
DE cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutain.
XX Hepatitis C virus.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 882 /note= Arg, Lys
FT Misc-difference 1184 /note= "Wild type Thr substituted by Ala"
FT
FT
XX WO200252015-A2.
XX 04-JUL-2002.
XX 20-DEC-2001; 2001WO-CA001843.
XX 22-DEC-2000; 2000US-0257857P.
XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX Kukolj G, Pause A;
XX WPI; 2002-575382/61.
XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.
XX
XX Claim 3; Page; 140pp; English.
XX The invention describes a self-replicating hepatitis C virus (HCV)
CC polynucleotide molecule comprising a 5'-non translated region (NTR),
CC where guanine at position 1 is substituted for adenine, a HCV polyprotein
CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at
CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(242)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence represents a mutant of the hepatitis
CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
CC This sequence does not appear in the specification but has been created
CC from the wild type sequence shown in ABG30580 using information given in
CC the claims of the invention
XX
XX Sequence 2201 AA;
SQ
Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAYCL 29
Db 855 GGVLAALAAYCL 866
RESULT 374
ABG30590
ID ABG30590 standard; protein; 2201 AA.
XX
AC ABG30590;
XX

DT 21-OCT-2002 (first entry)
XX Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B mutant #2.
XX Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;
DE cell culture replication; NS2/3; NS3/4; NS3; NS5B; mutant; mutain.
XX Hepatitis C virus.
OS Synthetic.
XX Key Location/Qualifiers
FH Misc-difference 339 /note= "Wild type Ser substituted by Gly"
FT Misc-difference 882 /label= Arg, Lys
FT
XX WO200252015-A2.
XX 04-JUL-2002.
XX 20-DEC-2001; 2001WO-CA001843.
XX 22-DEC-2000; 2000US-0257857P.
XX (BOEH) BOEHRINGER INGELHEIM CANADA LTD.
XX Kukolj G, Pause A;
XX WPI; 2002-575382/61.
XX New self-replicating RNA molecules from Hepatitis C virus (HCV), which
PT possess enhanced transduction or replication efficiency, useful for
PT evaluating potential inhibitors of HCV replication.
XX
XX Claim 3; Page; 140pp; English.
XX The invention describes a self-replicating hepatitis C virus (HCV)
CC polynucleotide molecule comprising a 5'-non translated region (NTR),
CC where guanine at position 1 is substituted for adenine, a HCV polyprotein
CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-
CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating
CC potential inhibitors of HCV replication. The HCV RNA molecule is also
CC useful for efficiently establishing cell culture replication. The self-
CC replicating polynucleotide molecule contains a 5'-NTR, where G at
CC position 1 is substituted for A, and therefore provides an alternative to
CC existing systems comprising a self-replicating HCV RNA molecule that, in
CC conjunction with mutations in the HCV non-structural region, such as the
CC G(242)C/R mutations, transduces and/or replicates with greater
CC efficiency. This amino acid sequence represents a mutant of the hepatitis
CC C virus replicon APGK12 and contains the viral protease NS2/3, protease
CC complex NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B. Note:
CC This sequence does not appear in the specification but has been created
CC from the wild type sequence shown in ABG30580 using information given in
CC the claims of the invention
XX
XX Sequence 2201 AA;
SQ
Query Match 10.2%; Score 12; DB 5; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAYCL 29
Db 855 GGVLAALAAYCL 866
RESULT 375
ABG30583
ID ABG30583 standard; protein; 2201 AA.
XX
AC ABG30583;
XX
DT 21-OCT-2002 (first entry)

KW vaccine; viral; bacterial; parasitic infection; prion disease;
 XX neoplastic; toxin; HCV; polyprotein.
 OS Hepatitis C virus.
 XX WO2004002415-A2.
 PN 08-JAN-2004.
 XX PD
 XX 27-JUN-2003; 2003WO-US020322.
 PF 27-JUN-2002; 2002US-0392718P.
 XX PR
 XX (DAND) DANA FARBER CANCER INST INC.
 XX PA
 XX Ruprecht RM, Jiang S;
 PI WPI; 2004-082868/08.
 XX DR
 XX Modulating an immune response, useful for treating immune disorders, e.g.
 PT viral, bacterial and parasitic infections, prion diseases, or neoplastic
 PT diseases, administering to a subject an overlapping synthetic peptide
 PT formulation.
 XX
 XX Claim 13; SEQ ID NO 211; 175pp; English.
 XX
 CC The invention relates to a novel method for modulating an immune response
 CC comprising administering to a subject an overlapping synthetic peptide
 CC formulation (OSPP) which comprises a combination of single chain peptides
 CC corresponding to the amino acid sequence of a protein of interest. The
 CC method of the invention has immunostimulant, virucide, antibacterial,
 CC antiparasitic and cytostatic applications and may be useful during
 CC vaccine production and for treating immune disorders including viral,
 CC bacterial and parasitic infections, prion diseases, neoplastic diseases,
 CC as well as providing protection against toxins. The current sequence is
 CC that of the OSPP-related Hepatitis C virus (HCV) polyprotein of the
 CC invention.
 XX
 XX Sequence 2280 AA;
 SQ
 Query Match 10.2%; Score 12; DB 8; Length 2280;
 Best Local Similarity 100.0%; Pred. No. 0.042; Mismatches 0; Gaps 0;
 Matches 12; Conservative 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCL 29
 Db 1664 GGVLAALAAAYCL 1675
 |||||
 |||||
 RESULT 378
 AAY70064
 ID AAY70064 standard; protein; 2307 AA.
 XX AC AAY70064;
 XX 12-SEP-2003 (revised)
 DT 05-JUN-2000 (first entry)
 XX
 XX Recombinant fusion pHCAP-1 polyprotein.
 DE
 XX Recombinant plasmid; pHCAP; Hepatitis C virus; HCV; reporter gene;
 KW NS3 protease; inhibitor; recombinant viral vector; RVV; HCV infection;
 KW secreted alkaline phosphatase; SEAP; serine protease; treatment;
 KW recombinant vaccinia virus.
 XX
 OS Hepatitis C virus.
 OS Vaccinia virus.
 OS Enterobacteria phage T7.
 OS Homo sapiens.
 OS Chimeric.
 XX
 Key Location/Qualifiers
 FH Region 1..92
 FT

FT Region
 FT 93..1784
 FT /label= pHCAP-1_polyprotein_fragment_1
 FT 93..98
 FT /label= pHCAP-1_polyprotein_fragment_2
 FT /note= "Additional residues resulting from subcloning of
 FT HCV/SEAP fragment"
 FT 100..390
 FT /label= E2/NS2 domain.
 FT /note= "Hepatitis C virus non-structural domain"
 FT 391..1028
 FT /label= NS3 domain
 FT /note= "Hepatitis C virus non-structural domain
 FT containing serine protease and helicase enzymes"
 FT 1019..1038
 FT /label= NS3/NS4A_cleavage site
 FT 1029..1082
 FT /label= NS4A domain
 FT /note= "Hepatitis C virus non-structural domain"
 FT 1076..1092
 FT /label= NS4A/NS4B_cleavage site
 FT 1083..1257
 FT /label= NS4B domain
 FT /note= "Hepatitis C virus non-structural domain"
 FT 1259..1278
 FT /label= NS5A/NS5B_cleavage site
 FT 1282..1784
 FT /label= Secreted alkaline phosphatase protein
 FT /note= "Secreted human placental SEAP"
 FT 1785..1936
 FT /label= pHCAP-1_polyprotein_fragment_3
 FT 1937..2021
 FT /label= pHCAP-1_polyprotein_fragment_4
 FT 2022..2307
 FT /label= pHCAP-1_polyprotein_fragment_5
 PN WO200008469-A1.
 XX
 XX 17-FEB-2000.
 XX
 XX 02-AUG-1999; 99WO-US017440.
 XX
 PR 05-AUG-1998; 98US-00129611.
 PR 08-MAR-1999; 99US-00263933.
 XX
 XX (AGOU-) AGOURON PHARM INC.
 XX
 XX Potts KE, Jackson RL, Patick AK;
 XX
 DR WPI; 2000-224057/19.
 DR N-PSDB; AAZ51002.
 XX
 XX Assessing compounds which augment or inhibit Hepatitis C virus NS3
 PT protease, useful particularly for identifying inhibitors which can be
 PT used for treating Hepatitis C virus infections.
 XX
 XX Claim 39; Page 70-78; 153pp; English.
 XX
 CC The patent discloses a reporter gene system for use in a cell-based
 CC assessment of inhibitors of Hepatitis C virus (HCV) NS3 protease, using a
 CC recombinant viral vector (RVV), that expresses a secreted human placental
 CC alkaline phosphatase (SEAP) reporter gene polyprotein, under the control
 CC of bacteriophage T7 promoter. The viral vector has been engineered to
 CC express a polyprotein, that includes NS3 HCV serine protease and the
 CC human SEAP gene. This assay system is useful for in vitro screening of
 CC potential protease inhibitors useful in the treatment of HCV infections
 CC and used to evaluate potent NS3 inhibitors, by monitoring the effect of
 CC increasing drug concentration on SEAP activity. NS3 inhibition is
 CC detected as a decrease in SEAP activity. The present sequence is the
 CC recombinant fusion protein, encoded by the RVV pHCAP-1, comprising five
 CC segments and the active NS2 and NS3 protease polyprotein, fused with the
 CC SEAP reporter protein. The plasmid is constructed using the pTM3 vector
 CC and has been used to generate recombinant vaccinia viruses. (Updated on
 CC 12-SEP-2003 to standardise OS field)

```
XX SQ Sequence 2307 AA;
Query Match 10.2%; Score 12; DB 3; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046

RESULT 379
AAY70065
ID AAY70065 standard; protein; 2307 AA.
XX AC AAY70065;
XX 12-SEP-2003 (revised)
DT 05-JUN-2000 (first entry)
XX Recombinant fusion pHCAP-3 polyprotein.
XX Recombinant plasmid; pHCAP; Hepatitis C virus; HCV; reporter gene;
KW NS3 protease; inhibitor; recombinant viral vector; RVV; HCV infection;
KW secreted alkaline phosphatase; SEAP; serine protease; treatment;
KW recombinant vaccinia virus.
XX Hepatitis C virus.
OS Vaccinia virus.
OS Enterobacteria phage T7.
OS Homo sapiens.
OS Chimeric.
PH Key Location/Qualifiers
FT Region 1..92
FT /label= pHCAP-1_polyprotein_fragment_1
FT Region 93..1784
FT /label= pHCAP-1_polyprotein_fragment_2
FT Region 93..98
FT /note= "Additional residues resulting from subcloning of
FT HCV/SEAP fragment"
FT Domain 100..390
FT /label= E2/NS2 domain
FT /note= "Hepatitis C virus non-structural domain"
FT Domain 391..1028
FT /label= NS3 domain
FT /note= "Hepatitis C virus non-structural domain
FT containing serine protease and helicase enzymes"
FT Misc-difference 536
FT /note= "Wild type catalytic Ser replaced with Ala, to
FT inactivate NS3 protease"
FT Cleavage-site 1019..1038
FT /label= NS3/NS4A_cleavage site
FT Domain 1029..1082
FT /label= NS4A domain
FT /note= "Hepatitis C virus non-structural domain"
FT Cleavage-site 1076..1092
FT /label= NS4A/NS4B_cleavage site
FT Domain 1083..1257
FT /note= "NS4B domain
FT /note= "Hepatitis C virus non-structural domain"
FT Cleavage-site 1259..1278
FT /label= NS5A/NS5B_cleavage site
FT Protein 1282..1784
FT /label= Secreted alkaline phosphatase protein
FT /note= "Secreted human placental SEAP"
FT Region 1785..1936
FT /label= pHCAP-1_polyprotein_fragment_3
FT Region 1937..2021
FT /label= pHCAP-1_polyprotein_fragment_4
FT Region 2022..2307
FT /label= pHCAP-1_polyprotein_fragment_5
```

```
XX PN WO200008469-A1.
XX 17-FEB-2000.
XX 02-AUG-1999; 99WO-US017440.
XX 05-AUG-1998; 98US-00129611.
XX 08-MAR-1999; 99US-00263933.
XX (AGOU-) AGOURON PHARM INC.
XX Potts KE, Jackson RL, Patick AK;
XX WPI; 2000-224057/19.
XX DR N-PSDB; AAZ51003.
XX Assessing compounds which augment or inhibit Hepatitis C virus NS3
XX protease, useful particularly for identifying inhibitors which can be
XX used for treating Hepatitis C virus infections.
XX Claim 40; Page 101-108; 153pp; English.
XX The patent discloses a reporter gene system for use in a cell-based
XX assessment of inhibitors of Hepatitis C virus (HCV) NS3 protease, using a
XX recombinant viral vector (RVV), that expresses a secreted human placental
XX alkaline phosphatase (SEAP) reporter gene polyprotein, under the control
XX of bacteriophage T7 promoter. The viral vector has been engineered to
XX express a polyprotein, that includes NS3 HCV serine protease and the
XX human SEAP gene. This assay system is useful for in vitro screening of
XX potential protease inhibitors useful in the treatment of HCV infections
XX and used to evaluate potent NS3 inhibitors, by monitoring the effect of
XX increasing drug concentration on SEAP activity. NS3 inhibition is
XX detected as a decrease in SEAP activity. The present sequence is the
XX recombinant fusion protein, encoded by the RVV pHCAP-3, comprising five
XX segments and the active NS2 and mutant NS3 protease, fused with the SEAP
XX reporter protein. The NS3 protease is inactivated by site directed
XX mutagenesis. The plasmid is constructed using the pTM3 vector and has
XX been used to generate recombinant vaccinia viruses. (Updated on 12-SEP-
XX 2003 to standardise OS field)
XX SQ Sequence 2307 AA;
Query Match 10.2%; Score 12; DB 3; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046

RESULT 380
AAY70066
ID AAY70066 standard; protein; 2307 AA.
XX AC AAY70066;
XX 12-SEP-2003 (revised)
DT 05-JUN-2000 (first entry)
XX Recombinant fusion pHCAP-4 polyprotein.
DE Recombinant plasmid; pHCAP; Hepatitis C virus; HCV; reporter gene;
KW NS3 protease; inhibitor; recombinant viral vector; RVV; HCV infection;
KW secreted alkaline phosphatase; SEAP; serine protease; treatment;
KW recombinant vaccinia virus.
XX Hepatitis C virus.
OS Vaccinia virus.
OS Enterobacteria phage T7.
OS Homo sapiens.
OS Chimeric.
```

XX FH Key Location/Qualifiers
FT Region 1..92
FT /label= pHCAP-1_polyprotein_fragment_1
FT Region 93..1784
FT /label= pHCAP-1_polyprotein_fragment_2
FT Region 93..98
FT /note= "Additional residues resulting from subcloning of
FT HCV/SEAP fragment"
FT Domain 100..390
FT /label= E2/NS2 domain
FT /note= "Hepatitis C virus non-structural domain"
FT Misc-difference 364
FT /note= "Wild type catalytic Cys replaced with Ala, to
FT inactivate NS2 protease"
FT Domain 391..1028
FT /label= NS3 domain
FT /note= "Hepatitis C virus non-structural domain
FT containing serine protease and helicase enzymes"
FT Misc-difference 536
FT /note= "Wild type catalytic Ser replaced with Ala, to
FT inactivate NS3 protease"
FT Cleavage-site 1019..1038
FT /label= NS3/NS4A_cleavage site
FT Domain 1029..1082
FT /label= NS4A domain
FT /note= "Hepatitis C virus non-structural domain"
FT Cleavage-site 1076..1092
FT /label= NS4A/NS4B_cleavage site
FT Domain 1083..1257
FT /label= NS4B domain
FT /note= "Hepatitis C virus non-structural domain"
FT Cleavage-site 1259..1278
FT /label= NS5A/NS5B_cleavage site
FT Protein 1282..1784
FT /label= "Secreted alkaline phosphatase protein
FT /note= "Secreted human placental SEAP"
FT Region 1785..1936
FT /label= pHCAP-1_polyprotein_fragment_3
FT Region 1937..2021
FT /label= pHCAP-1_polyprotein_fragment_4
FT Region 2022..2307
FT /label= pHCAP-1_polyprotein_fragment_5
XX WO200008469-A1.
XX 17-FEB-2000.
XX 02-AUG-1999; 99WO-US017440.
XX 05-AUG-1998; 98US-00129611.
XX 08-MAR-1999; 99US-00263933.
XX (AGOU-) AGOURON PHARM INC.
XX Potts KE, Jackson RL, Patick AK;
XX WPI; 2000-224057/19.
XX N-PSDB; AA251004.
XX Assessing compounds which augment or inhibit Hepatitis C virus NS3
XX protease, useful particularly for identifying inhibitors which can be
XX used for treating Hepatitis C virus infections.
XX Claim 41; Page 132-139; 153pp; English.
XX The patent discloses a reporter gene system for use in a cell-based
XX assessment of inhibitors of Hepatitis C virus (HCV) NS3 protease, using a
XX recombinant viral vector (RVV) that expresses a secreted human placental
XX alkaline phosphatase (SEAP) reporter gene polypeptide, under the control
XX of bacteriophage T7 promoter. The viral vector has been engineered to
XX express a polypeptide, that includes NS3 HCV serine protease and the
XX human SEAP gene. This assay system is useful for in vitro screening of

CC potential protease inhibitors useful in the treatment of HCV infections
CC and used to evaluate potent NS3 inhibitors, by monitoring the effect of
CC increasing drug concentration on SEAP activity. NS3 inhibition is
CC detected as a decrease in SEAP activity. The present sequence is the
CC recombinant fusion protein, encoded by the RVV pHCAP-4, comprising five
CC segments and the mutant inactive NS2 and mutant inactive NS3 protease
CC polyproteins, fused with the SEAP reporter protein. The NS2-NS3 domains
CC are inactivated by site directed mutagenesis. The plasmid is constructed,
CC using the pTM3 vector and has been used to generate recombinant vaccinia
CC viruses. (Updated on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 2307 AA;
SQ
Query Match 10.2%; Score 12; DB 3; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.043;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046
|||||
RESULT 381
AAR41435
ID AAR41435 standard; protein; 2354 AA.
XX AC AAR41435;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 24-FEB-1994 (first entry)
XX PT-NANBH virus non-structural proteins.
DE
KW Parenterally transmitted non A non B hepatitis; PT-NANBH; NS4;
KW Hepatitis C virus; HCV; detection; diagnosis; antigen; vaccine;
KW amplification; primer; polymerase chain reaction; PCR.
XX Hepatitis virus.
XX WO9317110-A2.
XX 02-SEP-1993.
XX 19-FEB-1993; 93WO-GB000345.
XX 21-FEB-1992; 92GB-00003803.
XX (WELL) WELLCOME FOUND LTD.
XX Parker D, Rodgers BC;
XX WPI; 1993-288415/36.
XX N-PSDB; AAQ46195.
XX New recombinant polypeptide for diagnosing hepatitis C - contains three
XX distinct antigens from different viral regions, also useful in protective
XX vaccines.
XX Example 1; Page 43-53; 99pp; English.
XX The NS4 region from the 3' region of the PT-NANBH genome (AAQ46195) is
XX amplified by PCR using primers D224 and D226 (AAQ46196-97) and the
XX fragment (AAQ46198) is cloned into a vector and expressed in infected
XX insect cells. The recombinant virus (BHC-19) was able to express the NS4
XX specific recombinant protein at low levels in the infected insect cells.
XX If at least three different PT-NANBH antigens are used to screen for PT-
XX NANBH, the screening is much more sensitive as compared to the use of
XX only two PT-NANBH antigens. Pref. antigens are described in AAQ46192-94.
XX Two new antigenic regions of the PT-NANBH genome are given in AAQ46198-
XX 99. AAQ46202 describes an improved PT-NANBH recombinant polypeptide.
XX (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to
XX correct OS field.)

```
XX SQ Sequence 2354 AA;
Query Match 10.2%; Score 12; DB 2; Length 2354;
Best Local Similarity 100.0%; Pred.No. 0.043;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 848 GGVLAALAAAYCL 859

RESULT 382
AAR29527
ID AAR29527 standard; protein; 2510 AA.
XX AC AAR29527;
XX OS Hepatitis C virus;
DT 25-MAR-2003 (revised)
DT 26-APR-1993 (first entry)
XX DE HCV antigen T7N1-30.
XX KW Clone; Hepatitis C Virus; HCV; core-envelope; NS1(gp70); NS2-NS4;
XX KW NS4-NS5; region; diagnostic method; antibody; suppress; control;
XX KW proteolytic; process; precursor; polypeptide.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
FT Misc-difference 2212 /note= "Nonsense codon"
XX FT
XX PN EP518313-A2.
XX PD 16-DEC-1992.
XX PF 11-JUN-1992; 92EP-00109812.
XX PR 11-JUN-1991; 91JP-00139268.
XX PR 12-JUL-1991; 91JP-00172794.
XX PR 07-OCT-1991; 91JP-00287008.
XX PR 16-DEC-1991; 91JP-00332329.
XX PR 20-APR-1992; 92JP-00099957.
XX PA (MITU ) MITSUBISHI KASEI CORP.
XX PI Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;
XX DR WPI; 1992-417213/51.
XX DR N-PSDB; AAQ32436.
XX PT New hepatitis C virus gene and its encoded protein - used for diagnosing
XX PT and vaccinating against hepatitis C virus infections.
XX PS Claim 1 and 3; Page 259-272; 305pp; English.
XX CC This sequence was encoded by the Hepatitis C Virus (HCV) gene of the
XX CC invention. The HCV gene is useful in the development of a diagnostic
XX CC method which is more accurate and effective than conventional ones; in
XX CC the detection of antibodies raised against a wide range of HCVs which
XX CC have been hardly detected before. The complete gene may be used in an
XX CC in vitro screening system for a substance capable of specifically suppressing
XX CC or controlling a proteolytic processing of a precursor polypeptide of
XX CC HCV. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 2510 AA;
Query Match 10.2%; Score 12; DB 2; Length 2510;
Best Local Similarity 100.0%; Pred.No. 0.046;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
```

```
Db 1664 GGVLAALAAAYCL 1675

RESULT 383
AEB17073
ID AEB17073 standard; protein; 2985 AA.
XX AC AEB17073;
XX DT 22-SEP-2005 (first entry)
XX DE Hepatitis C virus strain HCV-N polyprotein, SEQ ID NO: 40 #1.
XX KW RNA virus; microorganism; hepatitis C virus infection; antiinflammatory;
XX KW hepatotropic; virucide; gastrointestinal disease; infection; polyprotein.
XX OS Hepatitis C virus; Japanese genotype 1b (HCV-N).
XX FH Key Location/Qualifiers
FT Region 1..191 /note= "Core protein"
FT Region 192..383 /note= "Envelope protein (E1)"
FT Region 384..747 /note= "Envelope protein (E2)"
FT Region 748..810 /note= "P7 protein"
FT Region 811..1020 /note= "NS2 protein"
FT Misc-difference 1020..1021 /note= "Encoded by bases from position 3058 to 3153"
FT Region 1021..1628 /note= "NS3 protein"
FT Region 1629..1682 /note= "NS4A protein"
FT Region 1683..1943 /note= "NS4B protein"
FT Region 1944..2394 /note= "NS5A protein"
FT Misc-difference 2132 /note= "Encoded by CCA"
FT Misc-difference 2133 /note= "Encoded by TGC"
FT Misc-difference 2134 /note= "Encoded by GAG"
FT Misc-difference 2135 /note= "Encoded by CCC"
FT Misc-difference 2136 /note= "Encoded by GAA"
FT Misc-difference 2137 /note= "Encoded by CCG"
FT Misc-difference 2138 /note= "Encoded by GAC"
FT Misc-difference 2139 /note= "Encoded by GTA"
FT Misc-difference 2140 /note= "Encoded by GCA"
FT Misc-difference 2141 /note= "Encoded by GTG"
FT Misc-difference 2143 /note= "Encoded by ACT"
FT Misc-difference 2144 /note= "Encoded by TCC"
FT Misc-difference 2145 /note= "Encoded by ATG"
FT Misc-difference 2146 /note= "Encoded by CTC"
FT Misc-difference 2147 /note= "Encoded by ACC"
FT Misc-difference 2148 /note= "Encoded by GAC"
FT Misc-difference 2149
```


CC can be used as immunoassay reagents, for screening donated blood, and as
CC immunogens for vaccine prodn. Antibodies raised to the peptides can be
CC used in immunoassays to detect or quantify NANBV antigens in liver tissue
CC and blood. Preferred poly- peptides include residues 1-30, -115, or 2012;
CC 47-77; 116-191; 192-207 or -298; 230-238 or -263; 287-300; 293-330; 390-
CC 729; 730-1005; 1006-1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
CC and 2013-3010. The sequence is also disclosed in EP-464287 (SEQ ID NO 1).
CC See AAR20091 for details of this specification. (Updated on 25-MAR-2003
CC to correct PD field.) (Updated on 25-MAR-2003 to correct PA field.)
XX

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 385
AAR20091
ID AAR20091 standard; protein; 3010 AA.
XX
AC AAR20091;
XX
DT 25-MAR-2003 (revised)
DT 01-MAY-1992 (first entry)
XX
XX Non-A, non-B viral genome product.
XX
XX NANBV; vaccine; immunodiagnosis; antigen; antibody.
XX
XX Non-A.
OS non-B hepatitis virus.
OS
XX
XX Key Location/Qualifiers
FT Protein 1..115
FT /label= C
FT /note= "core protein"
FT 116..191
FT /label= M
FT /note= "matrix protein"
FT 192..389
FT /label= E
FT /note= "envelope protein"
FT 390..729
FT /label= NS1
FT Protein 730..1006
FT /label= NS2
FT 1007..1614
FT /label= NS3
FT Protein 1615..1862
FT /label= NS4a
FT Protein 1863..2012
FT /label= NS4b
FT Protein 2013..3010
FT /label= NS5
XX
XX EP464287-A.
PN
XX
XX 08-JAN-1992.
PD
XX
XX 28-DEC-1990; 90EP-00314371.
XX
XX 25-JUN-1990; 90JP-00167466.
XX 31-AUG-1990; 90JP-00230921.
PR 09-NOV-1990; 90JP-00305605.
PR 17-JUN-1991; 91EP-00401604.
XX
XX (OSAU) UNIV OSAKA.

DR WPI; 1992-009617/02.
DR N-PSDB; AAQ21829.
XX
PT New DNA from non-A, non-B hepatitis virus - and derived antigenic
PT polypeptide(s) useful for diagnostics, blood screening and in vaccines.
XX
PS Claim 3; Fig 2; 89pp; English.
XX
CC The sequence was deduced from several overlapping "BK" cDNA clones obtd.
CC by "gene walking" using a cDNA clone isolated from a library prepd. from
CC NANBV RNA. Antigenic polypeptides from the sequence can be used as
CC immunoassay reagents, for screening donated blood, and as immunogens for
CC vaccine prodn. Antibodies raised to the peptides can be used in
CC immunoassays to detect or quantify NANBV antigens in liver tissue and
CC blood. Preferred polypeptides are include residues 1-30, -115, or 2012;
CC 47-77; 116-191; 192-207 or -298; 230-238 or -263; 287-300; 293-330; 390-
CC 729; 730-1005; 1006-1614; 1384-1414; 1615-1862; 1737-1767; 1863-2012;
CC and 2013-3010. The sequence is also disclosed in EP-463848 (SEQ ID NO 1)
CC in which a virus particle contg. antigens encoded by the sequence is
CC claimed. See AAR20111 for details of this specification. (Updated on 25-
CC MAR-2003 to correct PA field.)
XX
XX Sequence 3010 AA;
SQ

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 386
AAR34580
ID AAR34580 standard; protein; 3010 AA.
XX
AC AAR34580;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 25-AUG-1993 (first entry)
XX
DE Human hepatitis C virus gene encoded polypeptide.
XX HCV; detection; diagnosis; vaccine; peptide.
XX
XX Hepatitis C virus.
OS
XX EP541089-A2.
PN
XX 12-MAY-1993.
PD
XX 05-NOV-1992; 92EP-00118974.
PF
XX 07-NOV-1991; 91JP-00318679.
PR
XX (SANW) SANWA KAGAKU KENKYUSHO CO.
PA
XX
XX Kuroono M, Mitani T, Jomori T, Hayashi Y, Suzuki E, Sawai K;
PI
XX
XX WPI; 1993-154074/19.
DR
XX N-PSDB; AAQ41345.
XX
XX Single or double stranded deoxyribonucleic acid for hepatitis C virus
XX detection - comprises 9500 nucleotide(s) and encodes human hepatitis C
XX virus gene, for diagnosis by polymerase chain reacting serum sample RNA
XX deriv. and electrophoresis.
XX
XX Disclosure; Page 4-20; 21pp; English.
XX
XX The sequence is that encoded by a human hepatitis C virus (HCV) gene. A
XX peptide comprising part of the amino acid sequence may be used as an

CC antigen as part of a method of detection for human anti-HCV antibody or
 CC in diagnosis of human hepatitis C. It may also be used as an antigen in a
 CC method of prodn. of vaccine for human hepatitis C. (Updated on 25-MAR-
 CC 2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||

Db 1664 GGVLAALAAAYCL 1675

RESULT 387

ID AAR30616 standard; protein; 3010 AA.

XX AAR30616;

27-AUG-2003 (revised)

25-MAR-2003 (revised)

19-MAY-1993 (first entry)

Polypeptide coded by Korean HCV full cDNA sequence LBC1.

KHCV-LBC1; diagnosis; vaccine.

Hepatitis C virus.

EP521318-A2.

07-JAN-1993.

10-JUN-1992; 92EP-00109753.

10-JUN-1991; 91KR-00009510.

06-AUG-1991; 91KR-00013601.

(LUCK-) LUCKY LTD.

Cho JM, Lee YB, Park YW, Lim KJ, Choi DY, So HS, Kim CH;

Kim ST, Yang JY;

WPI; 1993-001883/01.

N-PSDB; AAQ33282.

DNA and polypeptide(s) from a new type of hepatitis C virus (KHCV) - for
 diagnosing and vaccinating against KHCV infections.

Disclosure; Fig 2; 119pp; English.

The polypeptide is that encoded by the full cDNA sequence of Korean
 hepatitis C virus (KHCV) cDNA, KHCV-LBC1. It or its fragments may be used
 in a specific and accurate method for detecting KHCV antibodies in the
 serum of hepatitis C patients. Antibodies directed against these
 polypeptides are useful for the purification of KHCV antigens and for the
 development of an improved diagnostic to detect KHCV antigens in a
 sample. The polypeptides may also be used in a vaccine for treatment and
 prevention of KHCV infection at a dosage of 5-200 ug/peptide. (Updated on
 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
 field.)

SQ Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||

Db 1664 GGVLAALAAAYCL 1675

RESULT 388

ID AAR53417 standard; protein; 3010 AA.

XX AAR53417;

17-JAN-1995 (first entry)

Blood transmissible NANBHV protein.

Polymerase chain reaction; PCR; amplify; primer; non-A, non-B hepatitis;
 NANBH; virus; blood transmissible; detection; hepatitis virus; RT-PCR;
 C100 antibody; HCV RNA; NS5 region.

Non-A.
 non-B hepatitis virus.

Key Location/Qualifiers

Misc-difference 222 /label= His, Arg

Misc-difference 226 /label= Cys, Arg

Misc-difference 246 /label= Leu, Phe

Misc-difference 263 /label= Asp, Asn

Misc-difference 291 /label= Phe, Ser

Misc-difference 311 /label= Gly, Asp

Misc-difference 398 /label= Ser, Arg, Gly

Misc-difference 400 /label= Thr, Ala

Misc-difference 405 /label= Gln, Pro, Leu

Misc-difference 410 /label= Lys, Arg

Misc-difference 418 /label= Gly, Asp

Misc-difference 430 /label= Asn, Asp

Misc-difference 438 /label= Phe, Leu

Misc-difference 478 /label= Arg, Lys

Misc-difference 759 /label= Leu, Val

Misc-difference 1017 /label= Ser, Asn

Misc-difference 1036 /label= Thr, Ala

Misc-difference 1056 /label= Glu, Asp

Misc-difference 1201 /label= Met, Thr

Misc-difference 1205 /label= Met, Ile

Misc-difference 1255 /label= Asn, Tyr

Misc-difference 1263 /label= Gly, Asp

Misc-difference 1455 /label= Asn, Asp

Misc-difference 1828 /label= Ala, Thr

Misc-difference 1895 /label= Gly, Arg

Misc-difference 1896 /label= Gly, Ile

FT Region 417. .419
 /label= N-linked glycosylation site
 FT Region 423. .425
 /label= N-linked glycosylation site
 FT Region 430. .432
 /label= N-linked glycosylation site
 FT Region 448. .450
 /label= N-linked glycosylation site
 FT Region 532. .534
 /label= N-linked glycosylation site
 FT Region 556. .558
 /label= N-linked glycosylation site
 FT Region 576. .578
 /label= N-linked glycosylation site
 FT Region 623. .625
 /label= N-linked glycosylation site
 FT Region 645. .647
 /label= N-linked glycosylation site
 FT Region 1213. .1215
 /label= N-linked glycosylation site
 FT Region 1255. .1257
 /label= N-linked glycosylation site
 FT Region 2041. .2043
 /label= N-linked glycosylation site
 FT Region 2077. .2079
 /label= N-linked glycosylation site
 FT Region 2240. .2242
 /label= N-linked glycosylation site
 FT Region 2788. .2790
 /label= N-linked glycosylation site

JP06319583-A.

22-NOV-1994.

18-SEP-1992; 92JP-00249241.

18-SEP-1992; 92JP-00249241.

(SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.

WPI; 1995-040330/06.

N-PSDB; AAQ81559.

of hepatitis C virus helicase gene in baculovirus - useful for large scale prodn. of RNA helicase.

Claim 1; Fig 1-4; 9pp; Japanese.

AAQ81559 encodes AAR6864 hepatitis C virus (HCV) RNA helicase. The DNA was used in the construction of an expression vector, which was used to transform a baculovirus host. The transformed baculovirus could then be used for the recombinant prodn. of HCV RNA helicase

Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 2; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29

|||||

Db 1664 GGVLAAALAAAYCL 1675

RESULT 391

AAR82694

ID AAR82694 standard; protein; 3010 AA.

AC AAR82694;

DT 16-OCT-2003 (revised)

DT 14-NOV-1996 (first entry)

XX Partial HCV non-structural polyprotein.
 DE
 XX
 KW proteinase; hepatitis C virus; screening; inhibitor; proteolytic;
 KW identification; cleavage.
 XX
 OS Hepatitis C virus; Virus.
 XX
 PH Key Location/Qualifiers
 FT Protein 898. .1233
 FT /note= "partial proteinase; see AAR82692"
 FT Protein 992. .1907
 FT /note= "partial proteinase; see AAR82693"

XX JP07184648-A.

XX 25-JUL-1995.

XX 05-FEB-1993; 93JP-00018854.

XX 07-FEB-1992; 92JP-00022657.

XX 18-SEP-1992; 92JP-00249240.

XX 04-DEC-1992; 92JP-00325303.

XX (KAEN/) KAENNO K.

XX (SUMQ) SUMITOMO METAL IND LTD.

XX (SOYA-) SOYAKU GIJUTSU KENKYUSHO KK.

XX WPI; 1995-287962/38.

XX N-PSDB; AAT03960.

XX An HCV proteinase active substance - which has activity as an anti-HCV agent and can be used to screen for proteinase inhibitors.

XX Disclosure; Page 39-48; 52pp; Japanese.

XX The present sequence is a partial Hepatitis C Virus (HCV) polyprotein from the non-structural region. Partial proteinase sequences (AAR82692-93) are contained within this sequence. The proteinases can be used as anti-HCV agents. They can also be used to screen cpds. for their ability to inhibit their proteolytic activity. In this way proteinase inhibitors can be identified. (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 2; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29

|||||

Db 1664 GGVLAAALAAAYCL 1675

RESULT 392

AA06423

ID AA06423 standard; protein; 3010 AA.

AC AA06423;

DT 20-MAR-2003 (revised)

DT 27-SEP-1999 (first entry)

DE Non-A, non-B hepatitis virus polypeptide.

XX Non-A, non-B hepatitis B virus; NANBV; antigen; infection; diagnosis; vaccine.

XX Non-A.

OS non-B hepatitis virus.

XX Key Location/Qualifiers

FT Protein 1. .115

```

FT FT Protein /note= "core protein"
FT FT 116..191
FT FT Protein /note= "matrix protein"
FT FT 192..389
FT FT Protein /note= "envelope protein"
FT FT 390..729
FT FT Protein /note= "NS1 protein"
FT FT 730..1006
FT FT Protein /note= "NS2 protein"
FT FT 1007..1615
FT FT Protein /note= "NS3 protein"
FT FT 1616..1862
FT FT Protein /note= "NS4a protein"
FT FT 1863..2013
FT FT Protein /note= "NS4b protein"
FT FT 2014..3010
FT FT Protein /note= "NS5 protein"
XX EP933426-A1.
PN 04-AUG-1999.
XX 28-DEC-1990; 99EP-00106005.
XX 25-JUN-1990; 90JP-00167466.
PR 31-AUG-1990; 90JP-00303921.
PR 09-NOV-1990; 90JP-00305605.
PR 28-DEC-1990; 90EP-00314371.
XX (OSAU ) UNIV OSAKA.
XX Okayama H, Fuke I, Mori C, Takamizawa A, Yoshida I;
PI WPI; 1999-407152/35.
DR N-PSDB; AAX59394.
XX New hepatitis virus polypeptides, useful for diagnosing and treating
FT hepatitis infections.
XX Claim 2; Fig 2(1)-(16); 56pp; English.
XX This sequence represents the non-A, non-B hepatitis virus (NANBV)
CC polypeptide, as predicted from cDNA (see AAX59394) containing the entire
CC open reading frame of the NANBV genome. To obtain this cDNA, NANBV RNAs
CC were extracted directly from NANBV particles contained in whole blood of
CC a patient having NANB hepatitis, or from a resected liver of a patient
CC having NANB hepatitis and liver cancer. The RNA was then converted to
CC double-stranded cDNA. A cDNA library was produced and screened using
CC serum from a convalescent patient having acute NANB hepatitis and serum
CC from a patient having chronic NANB hepatitis. The isolated cDNA allows
CC recombinant production of NANBV antigen polypeptides in microbial or
CC eukaryotic cell culture. The method provides the safe production of NANBV
CC antigens with high purity on a large scale at low cost without the
CC biohazard associated with multiplying virus in animals. Claimed NANBV
CC nucleotide sequences are useful for the recombinant production of
CC polypeptides useful as antigens for vaccines, and as diagnostic reagents.
CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
CC correct PR field.)
XX Sequence 3010 AA;
SQ Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 393
AAX59022
ID AAX59022 standard; protein; 3010 AA.

/note= "core protein"
116..191
/note= "matrix protein"
192..389
/note= "envelope protein"
390..729
/note= "NS1 protein"
730..1006
/note= "NS2 protein"
1007..1615
/note= "NS3 protein"
1616..1862
/note= "NS4a protein"
1863..2013
/note= "NS4b protein"
2014..3010
/note= "NS5 protein"

EP933426-A1.
04-AUG-1999.
28-DEC-1990; 99EP-00106005.
25-JUN-1990; 90JP-00167466.
31-AUG-1990; 90JP-00303921.
09-NOV-1990; 90JP-00305605.
28-DEC-1990; 90EP-00314371.
(OSAU ) UNIV OSAKA.
Okayama H, Fuke I, Mori C, Takamizawa A, Yoshida I;
WPI; 1999-407152/35.
N-PSDB; AAX59394.
New hepatitis virus polypeptides, useful for diagnosing and treating
hepatitis infections.
Claim 2; Fig 2(1)-(16); 56pp; English.
This sequence represents the non-A, non-B hepatitis virus (NANBV)
polypeptide, as predicted from cDNA (see AAX59394) containing the entire
open reading frame of the NANBV genome. To obtain this cDNA, NANBV RNAs
were extracted directly from NANBV particles contained in whole blood of
a patient having NANB hepatitis, or from a resected liver of a patient
having NANB hepatitis and liver cancer. The RNA was then converted to
double-stranded cDNA. A cDNA library was produced and screened using
serum from a convalescent patient having acute NANB hepatitis and serum
from a patient having chronic NANB hepatitis. The isolated cDNA allows
recombinant production of NANBV antigen polypeptides in microbial or
eukaryotic cell culture. The method provides the safe production of NANBV
antigens with high purity on a large scale at low cost without the
biohazard associated with multiplying virus in animals. Claimed NANBV
nucleotide sequences are useful for the recombinant production of
polypeptides useful as antigens for vaccines, and as diagnostic reagents.
(Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
correct PR field.)
Sequence 3010 AA;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 393
AAX59022
ID AAX59022 standard; protein; 3010 AA.

```

```

XX AAX598022;
XX 21-JUN-1999 (first entry)
XX Infectious hepatitis C virus genotype 1b strain HC-J4 protein.
XX HCV; infectious clone; infection; diagnosis; therapy; vaccine; screening;
XX assay; antiviral; virucide.
XX Hepatitis C virus.
XX WO9904008-A2.
XX 28-JAN-1999.
XX 16-JUL-1998; 98WO-US014688.
XX 18-JUL-1997; 97US-0053062P.
XX 27-JAN-1998; 98US-00014416.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX WPI; 1999-132252/11.
XX N-PSDB; AAX24843.
XX New isolated hepatitis C virus nucleic acids - used to develop products
XX for the diagnosis, prevention and treatment of HCV infections and for
XX developing screening assays.
XX Claim 2; Fig 14G-H; 126pp; English.
XX This protein is encoded by the infectious hepatitis C virus (HCV)
XX genotype 1b strain HC-J4 genome (see AAX24843). HC-J4 was obtained from
XX acute phase plasma of a chimpanzee infected with serum containing HC-
XX J4/91. The infectious nucleic acid sequence can be used to produce
XX chimeric genomes (see AAX24833) consisting of the open reading frames of
XX infectious nucleic acid sequences of other genotypes (including genotypes
XX 1-6) and subtypes (such as 1b, 2a, 2b, 2c, 3a, 4a-f, 5a and 6a) of HCV.
XX The invention also relates to the introduction of mutations or deletions
XX into infectious nucleic acid sequences in order to produce an attenuated
XX HCV virus suitable for vaccine development. Infectious nucleic acid
XX sequences can also be used to produce attenuated virus via passage in
XX vitro or in vivo of the viruses produced by transfection of a host cell
XX with the infectious nucleic acid sequence. Vaccines comprising one or
XX more polypeptides made from the infectious nucleic acid sequence are used
XX to immunise mammals, especially humans, against hepatitis C. The nucleic
XX acid sequences can also be used to induce protective immunity against the
XX virus. The nucleic acid sequences or their encoded proteases (e.g. NS3
XX protease) can additionally be used to develop screening assays to
XX identify antiviral agents for HCV
XX Sequence 3010 AA;
SQ Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 394
AAB59174
ID AAB59174 standard; protein; 3010 AA.
XX AAB59174;
XX 21-MAR-2001 (first entry)
XX

```

DE Protein encoded by infectious Hepatitis C virus 1b genotype.
 XX GBV-B; hepatitis C virus; HCV; vaccine.
 KW Hepatitis C virus.
 OS WO200075337-A1.
 XX 14-DEC-2000.
 PD 02-JUN-2000; 2000WO-US015293.
 XX 04-JUN-1999; 99US-0137694P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Bukh J, Yanagi M, Emerson SU, Purcell RH;
 XX WPI; 2001-091214/10.
 DR New infectious nucleic acids of the GB virus-B clone, useful for
 PT indirectly studying the molecular properties of hepatitis C virus (HCV)
 PT and in developing vaccines and therapeutics for HCV.
 XX Disclosure; Fig 7; 96pp; English.
 PS The present invention relates to GB virus-B. The nucleic acid molecules
 XX of the invention are useful for indirectly studying the molecular
 CC properties of hepatitis C virus (HCV). The infectious nucleic acid
 CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
 CC in the development of vaccines and therapeutics for HCV
 XX Sequence 3010 AA;
 SQ

Query Match 10.2%; Score 12; DB 4; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 1664 GGVLAALAAAYCL 1675

RESULT 395
 AAB31170
 ID AAB31170 standard; protein; 3010 AA.
 AC AAB31170;
 XX 02-APR-2001 (first entry)
 DT Amino acid sequence of a hepatitis C virus (HCV) clone genotype 1b.
 DE Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
 XX HCV; vaccine; viral inhibitor; antiviral.
 KW Hepatitis C virus.
 OS WO200075352-A2.
 XX 14-DEC-2000.
 PD 02-JUN-2000; 2000WO-US015527.
 XX 04-JUN-1999; 99US-0137817P.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nam J, Bukh J, Emerson SU, Purcell RH;
 XX WPI; 2001-071081/08.
 DR N-P5DB; AAC86939.
 XX

PT New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
 PT in which the (non-)structural region has been replaced by hepatitis C
 PT virus (HCV) genome useful for treating or preventing HCV signs and
 XX symptoms.
 XX Disclosure; Fig 4G-H; 97pp; English.
 PS The specification describes a nucleic acid comprising a chimeric virus
 CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
 CC the (non-)structural region has been replaced by the (non-)structural
 CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
 CC the chimeric virus and the chimeric virus are useful for identifying cell
 CC lines capable of supporting the replication of these chimeric viruses, in
 CC screening for neutralizing antibodies to HCV of different genotypes, in
 CC the production of HCV-BVDV virions, for the development of inactivated or
 CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
 CC molecular properties of HCV indirectly in vitro, and in identifying
 CC inhibitors of viral enzyme activity which would be useful as antiviral
 CC agents. Formulations or compositions comprising the chimeric virions may
 CC be used to treat or prevent the signs and symptoms of HCV. The present
 CC sequence is encoded by a HCV clone, which is used to construct chimeric
 CC nucleic acids of the invention
 XX Sequence 3010 AA;
 SQ

Query Match 10.2%; Score 12; DB 4; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.053;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 1664 GGVLAALAAAYCL 1675

RESULT 396
 ABG32458
 ID ABG32458 standard; protein; 3010 AA.
 XX AC ABG32458;
 XX 15-NOV-2002 (first entry)
 DT Hepatitis C virus Con 1 isolate polyprotein mutant #7.
 XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
 XX Hepatitis C virus.
 OS Synthetic.
 XX Key Location/Qualifiers
 FT Misc-difference 2197 /note= "Wild-type Phe substituted by Ser"
 FT WO200259321-A2.
 XX 01-AUG-2002.
 PD 16-JAN-2002; 2002WO-EP000526.
 XX 23-JAN-2001; 2001US-0263479P.
 XX (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 PT

```
XX PS Claim 1; Page: 69pp; English.
XX CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins). NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX SQ Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 5; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 397
ABG32459
ID ABG32459 standard; protein; 3010 AA.
XX AC ABG32459;
XX DT 15-NOV-2002 (first entry)
XX DE Hepatitis C virus Con 1 isolate polyprotein mutant #8.
XX KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutein;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX Misc-difference 2198
XX FT /note= "Wild-type Leu substituted by Ser"
XX PN WO200259321-A2.
XX PD 01-AUG-2002.
XX PF 16-JAN-2002; 2002WO-EP000536.
XX PR 23-JAN-2001; 2001US-0263479P.
XX PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX PI De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.

XX PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX PS Claim 1; Page: 69pp; English.
XX CC The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins). NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX SQ Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 5; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 398
ABG32451
ID ABG32451 standard; protein; 3010 AA.
XX AC ABG32451;
XX DT 15-NOV-2002 (first entry)
XX DE Hepatitis C virus Con 1 isolate polyprotein.
XX KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX OS Hepatitis C virus.
XX PN WO200259321-A2.
XX PD 01-AUG-2002.
XX PF 16-JAN-2002; 2002WO-EP000526.
XX PR 23-JAN-2001; 2001US-0263479P.
XX PA (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX PI De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
```

DR N-PSDB; ABK91411.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 1; Page 34-36; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
XX polypeptide (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
XX and NS5B proteins) used as a basis for the adaptive mutations of the
XX invention
XX
XX Sequence 3010 AA;
Query Match 10.2%; Score 12; DB 5; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 399
ABG32455
ID ABG32455 standard; protein; 3010 AA.
XX AC ABG32455;
XX DT 15-NOV-2002 (first entry)
XX DE Hepatitis C virus Con 1 isolate polyprotein mutant #4.
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX Misc-difference 2041 /note= "Wild-type Asn substituted by Thr"
XX FT
XX WO200259321-A2.
XX PD 01-AUG-2002.
XX PF 16-JAN-2002; 2002WO-EP000526.
XX PR 23-JAN-2001; 2001US-0263479P.
XX PA (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.

XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 1; Page; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
XX polypeptide (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
XX and NS5B proteins), NS5A mutant of the invention. Note: The present
XX sequence is not shown in the specification but was created by the indexer
XX using the HCV sequence appearing as ABG32451 and the information in claim
XX 1
XX Sequence 3010 AA;
Query Match 10.2%; Score 12; DB 5; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 400
ABG32457
ID ABG32457 standard; protein; 3010 AA.
XX AC ABG32457;
XX DT 15-NOV-2002 (first entry)
XX DE Hepatitis C virus Con 1 isolate polyprotein mutant #6.
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX Misc-difference 2173 /note= "Wild-type Phe substituted by Ser"
XX FT
XX WO200259321-A2.
XX PD 01-AUG-2002.

```
PF 16-JAN-2002; 2002WO-EP000526.
XX
PR 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX
XX Claim 1; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX
XX Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 5; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
XX
XX RESULT 401
XX ABG32460
XX ID ABG32460 standard; protein; 3010 AA.
XX
XX AC ABG32460;
XX
XX DT 15-NOV-2002 (first entry)
XX
XX Hepatitis C virus Con 1 isolate polyprotein mutant #9.
XX
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutetin;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
XX Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 2199 /note= "Wild-type Ala substituted by Thr"
XX
```

```
XX
PN WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-EP000526.
XX
XX 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
PT ribosome entry site (IRES) region, useful in studying HCV replication and
PT expression.
XX
XX Claim 1; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
CC internal ribosome entry site (IRES) region coding for one or more NS3,
CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
CC are detailed in the specification. Also included are (1) an expression
CC vector comprising a nucleotide sequence coding for the altered nucleic
CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
CC recombinant cell human hepatoma cell comprising the altered nucleic acids
CC ; (3) a recombinant cell produced by introducing into a human hepatoma
CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
CC replicon enhanced cell or which containing a functional HCV replicon; (5)
CC an HCV replicon enhanced cells made in the method; and (6) measuring the
CC ability of a compound to affect HCV activity. The HCV replicons and HCV
CC replicon enhanced cells are useful in studying HCV replication and
CC expression, and HCV and host cell interactions, producing HCV RNA and
CC proteins, and providing a system for measuring the ability of a compound
CC to modulate one or more HCV activities e.g. to discover drugs which may
CC treat HCV mediated diseases such as liver failure, cirrhosis and
CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
CC and NS5B proteins), NS5A mutant of the invention. Note: The present
CC sequence is not shown in the specification but was created by the indexer
CC using the HCV sequence appearing as ABG32451 and the information in claim
CC 1
XX
XX Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 5; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
XX
XX RESULT 402
XX ABG32453
XX ID ABG32453 standard; protein; 3010 AA.
XX
XX AC ABG32453;
XX
XX DT 15-NOV-2002 (first entry)
XX
XX Hepatitis C virus Con 1 isolate polyprotein mutant #2.
XX
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutetin;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
XX Hepatitis C virus.
OS
```

OS Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 1202
 FT /note= "Wild-type Glu substituted by Gly"
 XX
 XX WO200259321-A2.
 XX
 XX 01-AUG-2002.
 XX
 XX 16-JAN-2002; 2002WO-EP000526.
 XX
 XX 23-JAN-2001; 2001US-0263479P.
 XX
 XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
 XX
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX
 XX Claim 1; Page; 69pp; English.
 XX
 XX The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV
 CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
 CC polypeptide (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
 CC and NS5B proteins), NS3 mutant of the invention. Note: The present
 CC sequence is not shown in the specification but was created by the indexer
 CC using the HCV sequence appearing as ABG32451 and the information in claim
 CC 1
 XX
 XX Sequence 3010 AA;
 XX
 XX Query Match 10.2%; Score 12; DB 5; Length 3010;
 XX Best Local Similarity 100.0%; Pred. No. 0.053;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 18 GGVLAALAAAYCL 29
 XX |||||
 XX Db 1664 GGVLAALAAAYCL 1675
 XX
 XX RESULT 403
 XX ABG32461
 XX ID ABG32461 standard; protein; 3010 AA.
 XX AC ABG32461;
 XX AC
 XX 15-NOV-2002 (first entry)
 XX DT
 XX Hepatitis C virus Con 1 isolate polypeptide mutant #10.
 XX DE
 XX

KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
 XX internal ribosome entry site; IRES; NS5A; HCV replication; polypeptide.
 OS Hepatitis C virus.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 2204
 FT /note= "Wild-type Ser substituted by Arg"
 XX
 XX WO200259321-A2.
 XX
 XX 01-AUG-2002.
 XX
 XX 16-JAN-2002; 2002WO-EP000526.
 XX
 XX 23-JAN-2001; 2001US-0263479P.
 XX
 XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
 XX
 XX De Francesco R, Migliaccio G, Paonessa G;
 XX WPI; 2002-599793/64.
 XX
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
 PT ribosome entry site (IRES) region, useful in studying HCV replication and
 PT expression.
 XX
 XX Claim 1; Page; 69pp; English.
 XX
 XX The invention relates to nucleic acid molecules comprising altered HCV
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
 CC internal ribosome entry site (IRES) region coding for one or more NS3,
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations
 CC are detailed in the specification. Also included are (1) an expression
 CC vector comprising a nucleotide sequence coding for the altered nucleic
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
 CC replicon enhanced cell or which containing a functional HCV replicon; (5)
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV
 CC replicon enhanced cells are useful in studying HCV replication and
 CC expression, and HCV and host cell interactions, producing HCV RNA and
 CC proteins, and providing a system for measuring the ability of a compound
 CC to modulate one or more HCV activities e.g. to discover drugs which may
 CC treat HCV mediated diseases such as liver failure, cirrhosis and
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
 CC polypeptide (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A
 CC and NS5B proteins), NS5A mutant of the invention. Note: The present
 CC sequence is not shown in the specification but was created by the indexer
 CC using the HCV sequence appearing as ABG32451 and the information in claim
 CC 1
 XX
 XX Sequence 3010 AA;
 XX
 XX Query Match 10.2%; Score 12; DB 5; Length 3010;
 XX Best Local Similarity 100.0%; Pred. No. 0.053;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 18 GGVLAALAAAYCL 29
 XX |||||
 XX Db 1664 GGVLAALAAAYCL 1675
 XX
 XX RESULT 404
 XX ABG32454
 XX ID ABG32454 standard; protein; 3010 AA.
 XX AC ABG32454;
 XX

XX 15-NOV-2002 (first entry)
DT Hepatitis C virus Con 1 isolate polyprotein mutant #3.
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutetin;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.
XX Hepatitis C virus.
OS Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 1347 /note= "Wild-type Ala substituted by Thr"
XX WO200259321-A2.
XX 01-AUG-2002.
XX 16-JAN-2002; 2002WO-EP000526.
XX 23-JAN-2001; 2001US-0263479P.
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 1; Page; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NSSA, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX ; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
XX and NSSB proteins), NS3 mutant of the invention. Note: The present
XX sequence is not shown in the specification but was created by the indexer
XX using the HCV sequence appearing as ABG32451 and the information in claim
XX 1
SQ Sequence 3010 AA;
Query Match 10.2%; Score 12; DB 5; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 405
ABG32452
ID ABG32452 standard; protein; 3010 AA.
XX AC ABG32452;
XX 15-NOV-2002 (first entry)
DT Hepatitis C virus Con 1 isolate polyprotein mutant #1.
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutetin;
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
XX internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.
XX Hepatitis C virus.
OS Synthetic.
XX Key Location/Qualifiers
XX Misc-difference 1095 /note= "Wild-type Gly substituted by Ala"
XX WO200259321-A2.
XX 01-AUG-2002.
XX 16-JAN-2002; 2002WO-EP000526.
XX 23-JAN-2001; 2001US-0263479P.
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 1; Page; 69pp; English.
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NSSA, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX ; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is the HCV replicon Con 1
XX and NSSB proteins), NS3 mutant of the invention. Note: The present
XX sequence is not shown in the specification but was created by the indexer
XX using the HCV sequence appearing as ABG32451 and the information in claim
XX 1
SQ Sequence 3010 AA;
Query Match 10.2%; Score 12; DB 5; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 406
ADP88597
ID ADF88597 standard; protein; 3010 AA.
XX AC ADF88597;
XX DT 26-FEB-2004 (first entry)
XX DE Hepatitis C virus NS3 gene protein, SEQ ID No 6.
XX KW liver cancer; hepatitis-C virus; NS3 gene; carcinogenicity.
XX OS Hepatitis C virus.
XX PN JP2003210181-A.
XX PD 29-JUL-2003.
XX PF 30-MAY-2002; 2002JP-00158335.
XX PR 16-NOV-2001; 2001JP-00352443.
XX PA (SHIN-) ZH SHINSANGYO SOZO KENKYU KIKO.
XX DR WPI; 2003-819836/77.
XX DR N-FSDB; ADF88596.
XX PT Diagnosing liver cancer, involves amplifying amino terminal region of hepatitis-C virus gene using predetermined primer and determining hepatitis-C virus in base sequence of obtained DNA fragment.
XX PS Disclosure; SEQ ID NO 6; 36pp; Japanese.
XX CC The invention relates to the novel testing method for diagnosing liver cancer. The novel method comprises amplifying the amino terminal region of a hepatitis-C virus NS3 gene using a predetermined primer and determining the hepatitis-C virus in a base sequence of the obtained DNA fragment. The novel testing method is useful for diagnosing liver cancer and also used in a gene amplification technique, a clinical laboratory test reagent, a polymerase chain reaction, a base sequence analysis and genetic engineering. The method enables the detection of a hepatitis-C virus having high carcinogenicity with high specificity. This sequence represents the protein of the hepatitis-C virus NS3 gene of the invention.

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

Query Match 10.2%; Score 12; DB 7; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 407
ADO36227
ID ADO36227 standard; protein; 3010 AA.
XX AC ADO36227;
XX DT 26-AUG-2004 (first entry)
XX DE Hepatitis C virus (HCV) J4L6 wild-type polyprotein.
XX KW hepatotropic; virucide; vaccine; gene therapy; vaccine;
XX Hepatitis C virus; HCV; core protein; HCV infection; vaccination;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

Query Match 10.2%; Score 12; DB 8; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 408
ADO79401
ID ADO79401 standard; protein; 3010 AA.
XX AC ADO79401;
XX DT 26-AUG-2004 (first entry)
XX DE Hepatitis C virus J4L6 genome wild-type polyprotein.
XX KW HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN WO2004046176-A1.
XX PD 03-JUN-2004.
XX PF 13-NOV-2003; 2003WO-EP012830.

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

Query Match 10.2%; Score 12; DB 8; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

polyprotein.
Hepatitis C virus.
WO2004046175-A1.
03-JUN-2004.
13-NOV-2003; 2003WO-EP012793.
15-NOV-2002; 2002GB-00026722.
(GLAX) GLAXO GROUP LTD.
Brett S, Hamblin PA, Ogilvie L;
WPI; 2004-420613/39.
N-FSDB; ADO36222.
New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes the polypeptide sequences of the HCV core and at least one other HCV protein, for use in medicine, particularly for manufacturing a medicament for treating HCV.
Disclosure; Fig 6; 78pp; English.
The invention describes a polynucleotide vaccine comprising a polynucleotide sequence (SI) encoding the Hepatitis C virus (HCV) Core protein and at least 1 other HCV protein, and causes expression of the proteins in cells (in which (SI) has been mutated or positioned relative to the polynucleotide sequence encoding the other HCV protein, so that the negative effect of the Core protein on expression of the other HCV protein is reduced). Also described are: a method of preventing or treating an HCV infection in a mammal, comprising administering the vaccine cited above to a mammal; and a method of vaccination of an individual, comprising taking a polynucleotide vaccine as cited above, coating the polynucleotide onto gold beads and delivering the gold beads into the skin. HCV nucleic acids, polypeptides, host cells, vectors and antibodies used in the methods, are also disclosed. The polynucleotide vaccine is useful in the manufacture of a medicament for the treatment of HCV. This is the amino acid sequence of the wild type HCV polyprotein.
Sequence 3010 AA;
Query Match 10.2%; Score 12; DB 8; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 408
ADO79401
ID ADO79401 standard; protein; 3010 AA.
XX AC ADO79401;
XX DT 26-AUG-2004 (first entry)
XX DE Hepatitis C virus J4L6 genome wild-type polyprotein.
XX KW HCV; polyprotein; vaccine; DNA immunisation; hepatotropic; virucide.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN WO2004046176-A1.
XX PD 03-JUN-2004.
XX PF 13-NOV-2003; 2003WO-EP012830.

XX PR 15-NOV-2002; 2002GB-00026722.
XX (GLAX) GLAXO GROUP LTD.
XX PI Brett S, Hamblin PA, Ogilvie L;
XX WPI; 2004-420614/39.
DR N-PSDB; ADO79396.
XX
XX New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
PT for use in medicine, in particular for manufacturing a medicament for the
PT treatment of HCV.
XX
XX Disclosure; Fig 6; 79pp; English.
XX
XX The present sequence is the hepatitis C virus (HCV) J4L6 genome wild-type
CC polypeptide sequence. HCV DNA vaccines of the invention comprise a
CC polynucleotide that encodes the HCV proteins Core, NS3, NS4B and NS5B,
CC and does not encode the NS4A and/or NS5A proteins. Preferably, the codon
CC usage of the polynucleotide sequence resembles that of highly expressed
CC human genes. The polynucleotides may encode individual proteins or fusion
CC proteins. Preferred fusions include double fusions between NS4B and NS5B
CC and between Core and NS3. The vaccines are useful for the treatment or
CC prevention of an HCV infection.
XX
XX Sequence 3010 AA;
SQ
Query Match 10.2%; Score 12; DB 8; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 409
ADX40800
ID ADX40800 standard; protein; 3010 AA.
AC ADX40800;
XX
XX 21-APR-2005 (first entry)
XX HCV polymerase protein #23.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis C virus.
OS
XX WO2005012502-A2.
XX
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX HCV polymerase protein #23.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis C virus.
OS
XX WO2005012502-A2.
XX
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprising identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
SQ
Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 410
ADX40801
ID ADX40801 standard; protein; 3010 AA.
XX
XX ADX40801;
AC
XX
XX 21-APR-2005 (first entry)
XX HCV polymerase protein #24.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis C virus.
OS
XX WO2005012502-A2.
XX
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprising identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
SQ
Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
SQ
Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 410
ADX40801
ID ADX40801 standard; protein; 3010 AA.
XX
XX ADX40801;
AC
XX
XX 21-APR-2005 (first entry)
XX HCV polymerase protein #24.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis C virus.
OS
XX WO2005012502-A2.
XX
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprising identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
SQ
Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 411
ADX40811
ID ADX40811 standard; protein; 3010 AA.
XX
XX
AC ADX40811;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #34.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
PN WO2005012502-A2.
XX
XX
PD 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCL 29
XX 1664 GGVLAALAAAYCL 1675
XX
XX RESULT 412
XX ADX40795
XX ID ADX40795 standard; protein; 3010 AA.
XX
XX AC ADX40795;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE HCV polymerase protein #18.
XX
XX KW Immune stimulation; polymerase; enzyme.
XX
XX OS Hepatitis C virus.
XX
XX PN WO2005012502-A2.
XX
XX

PD 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAAAYCL 29
XX 1664 GGVLAALAAAYCL 1675
XX
XX RESULT 413
XX ADX40796
XX ID ADX40796 standard; protein; 3010 AA.
XX
XX AC ADX40796;
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE HCV polymerase protein #19.
XX
XX KW Immune stimulation; polymerase; enzyme.
XX
XX OS Hepatitis C virus.
XX
XX PN WO2005012502-A2.
XX
XX PD 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.
XX
XX

XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PS Disclosure; Page 387-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675
RESULT 417
ADX40786
ID ADX40786 standard; protein; 3010 AA.
XX AC ADX40786;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #9.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PS Disclosure; Page 387-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675
RESULT 417
ADX40786
ID ADX40786 standard; protein; 3010 AA.
XX AC ADX40786;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #9.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX

PS Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675
RESULT 418
ADX40817
ID ADX40817 standard; protein; 3010 AA.
XX AC ADX40817;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #40.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PS Disclosure; Page 388-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 18 GGVLAALAAAYCL 29

```
Db      1664 GGVLAALAAAYCL 1675
|||||
RESULT 419
ADX40794
ID   ADX40794 standard; protein; 3010 AA.
XX AC ADX40794;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #17.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WI WI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
XX
XX Query Match      10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy      18 GGVLAALAAAYCL 29
Db      1664 GGVLAALAAAYCL 1675
|||||
RESULT 421
ADX40791
ID   ADX40791 standard; protein; 3010 AA.
XX AC ADX40791;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #14.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WI WI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
XX
XX Query Match      10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy      18 GGVLAALAAAYCL 29
Db      1664 GGVLAALAAAYCL 1675
|||||
RESULT 420
ADX40807
ID   ADX40807 standard; protein; 3010 AA.
XX AC ADX40807;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #30.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX
```

XX Disclosure; Page 388-440; 458pp; English.
XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;
SQ

Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 422
ADX40792
ID ADX40792 standard; protein; 3010 AA.

XX AC ADX40792;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #15.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX FN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 422
ADX40792
ID ADX40792 standard; protein; 3010 AA.

XX AC ADX40792;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #6.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 423
ADX40781
ID ADX40781 standard; protein; 3010 AA.

XX AC ADX40781;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #4.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX FN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PF 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX
XX Disclosure; Page 387-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX
XX Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 424
ADX40783
ID ADX40783 standard; protein; 3010 AA.

XX AC ADX40783;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #6.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.

XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PP 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PS Disclosure; Page 388-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX PT I binding motif.
XX QY 18 GGVLAALAAAYCL 29
XX DB 1664 GGVLAALAAAYCL 1675
XX RESULT 425
XX ADX40805
XX ID ADX40805 standard; protein; 3010 AA.
XX AC ADX40805;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #28.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PP 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class

PT I binding motif.
XX PS Disclosure; Page 388-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAALAAAYCL 29
XX DB 1664 GGVLAALAAAYCL 1675
XX RESULT 426
XX ADX40816
XX ID ADX40816 standard; protein; 3010 AA.
XX AC ADX40816;
XX DT 21-APR-2005 (first entry)
XX DE HCV polymerase protein #39.
XX KW Immune stimulation; polymerase; enzyme.
XX OS Hepatitis C virus.
XX PN WO2005012502-A2.
XX PD 10-FEB-2005.
XX PP 29-MAR-2004; 2004WO-US009510.
XX PR 28-MAR-2003; 2003US-0458026P.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX DR WPI; 2005-132661/14.
XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX PS Disclosure; Page 388-440; 458pp; English.
XX CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX SQ Sequence 3010 AA;
XX Query Match 10.2%; Score 12; DB 9; Length 3010;
XX Best Local Similarity 100.0%; Pred. No. 0.053;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 427
ADX40788
ID ADX40788 standard; protein; 3010 AA.
XX
AC ADX40788;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #11.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
FN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
SQ Sequence 3010 AA;
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 428
ADX40789
ID ADX40789 standard; protein; 3010 AA.
XX
AC ADX40789;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #12.
XX
KW Immune stimulation; polymerase; enzyme.
XX
FN
XX
PD
XX
PF
XX
PR
XX
PA
XX
PI
XX
DR
XX
PT

Query Match 10.2%; Score 12; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 429
ADX40804
ID ADX40804 standard; protein; 3010 AA.
XX
AC ADX40804;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #27.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
FN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
SQ Sequence 3010 AA;
XX
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.

PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.

PS Disclosure; Page 388-440; 458pp; English.

XX
XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.

XX Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 9; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

RESULT 430

ADX40806

ID ADX40806 standard; protein; 3010 AA.

AC ADX40806;

DT 21-APR-2005 (first entry)

DE HCV polymerase protein #29.

XX Immune stimulation; polymerase; enzyme.

XX Hepatitis C virus.

XX WO2005012502-A2.

XX 10-FEB-2005.

XX 29-MAR-2004; 2004WO-US009510.

XX 28-MAR-2003; 2003US-0458026P.

XX (EPIM-) EPIMMUNE INC.

XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

XX WPI; 2005-132661/14.

XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.

PS Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.

XX Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 9; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

RESULT 431

ADX40812

ID ADX40812 standard; protein; 3010 AA.

XX ADX40812;

XX 21-APR-2005 (first entry)

XX HCV polymerase protein #35.

XX Immune stimulation; polymerase; enzyme.

XX Hepatitis C virus.

XX WO2005012502-A2.

XX 10-FEB-2005.

XX 29-MAR-2004; 2004WO-US009510.

XX 28-MAR-2003; 2003US-0458026P.

XX (EPIM-) EPIMMUNE INC.

XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

XX WPI; 2005-132661/14.

XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.

PS Disclosure; Page 388-440; 458pp; English.

XX The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.

XX Sequence 3010 AA;

Query Match 10.2%; Score 12; DB 9; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

RESULT 432

AAR34468

ID AAR34468 standard; protein; 3011 AA.

XX AAR34468;

XX 30-JUL-1993 (first entry)

XX Encoded by full-length Hepatitis C virus clone JKI-B.

XX HCV; non-A, non-B hepatitis virus; NANBHV; liver disease;

polymerase chain reaction; diagnostic method.
Hepatitis C virus.
Key Location/Qualifiers
FT Misc-difference 2414
FT /note= "not defined"
XX
PN JP05068562-A.
XX
XX 23-MAR-1993.
XX
XX 30-MAY-1991; 91JP-00153736.
XX
XX 30-MAY-1991; 91JP-00153736.
XX
XX (SANW) SANWA KAGAKU KENKYUSHO CO.
XX
XX WPI; 1993-130638/16.
DR N-PSDB; AAQ40426.
XX
XX DNA and cDNA of hepatitis C virus - useful as probes for diagnosing HCV infection.
XX
XX Claim 3; Page 6-18; 44pp; Japanese.
XX
XX cDNA was prepared from HCV genomic RNA. Full-length clone JKI-B and 14 shorter clones were isolated by PCR amplification. Primer/probes derived from the sequences of these clones can be used in diagnostic assays for HCV. See also AAQ40425-Q40439
XX
XX Sequence 3011 AA;
Query Match 10.2%; Score 12; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 433
ABG32456
ID ABG32456 standard; protein; 3011 AA.
XX
AC ABG32456;
XX
DT 15-NOV-2002 (first entry)
XX
DE Hepatitis C virus Con 1 isolate polyprotein mutant #5.
XX
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;
KW Hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.
XX
XX Hepatitis C virus.
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 2039. .2041
FT /note= "Wild-type Lys-Asn substituted by Lys-Lys-Asn"
XX
XX WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-BF000526.
XX
XX 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX

De Francesco R, Migliaccio G, Paonessa G;
WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region, useful in studying HCV replication and expression.
XX
XX Claim 1; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region coding for one or more NS3, NS5A, or EMCV IRES mutations, respectively. The location of the mutations are detailed in the specification. Also included are (1) an expression vector comprising a nucleotide sequence coding for the altered nucleic acids, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus) replicon enhanced cell or which containing a functional HCV replicon; (5) an HCV replicon enhanced cells made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicons and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is the HCV replicon Con 1 polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A and NS5B proteins), NS5A mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the HCV sequence appearing as ABG32451 and the information in claim 1
XX
XX Sequence 3011 AA;
Query Match 10.2%; Score 12; DB 5; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 434
ADX40797
ID ADX40797 standard; protein; 3013 AA.
XX
AC ADX40797;
XX
DT 21-APR-2005 (first entry)
XX
XX HCV polymerase protein #20.
XX
XX Immune stimulation; polymerase; enzyme.
XX
XX Hepatitis C virus.
XX
XX WO2005012502-A2.
XX
XX 10-FEB-2005.
XX
XX 29-MAR-2004; 2004WO-US009510.
XX
XX 28-MAR-2003; 2003US-0458026P.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX

DR WPI; 2005-132661/14.

XX Identifying a candidate peptide epitope, which induces a HLA class I CTL

PT response comprises identifying variants of a peptide epitope 8-11 amino

PT acids in length comprising primary anchor residues of the same HLA class

PT I binding motif.

XX

PS Disclosure; Page 388-440; 458pp; English.

XX

CC The invention relates to a method of identifying a candidate peptide

CC epitope which induces an HLA class I CTL response against variants of the

CC peptide epitope, comprising identifying, from a particular antigen of an

CC infectious agent, variants of a peptide epitope comprising primary anchor

CC residues of the same HLA class I binding motif. The method is useful for

CC identifying a candidate peptide epitope, which induces an HLA class I CTL

CC response against variants of the peptide epitope. This sequence

CC represents an HCV polymerase protein used in the scope of the invention.

XX

XX Sequence 3013 AA;

SQ

Query Match 10.2%; Score 12; DB 9; Length 3013;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

RESULT 435

AAR35207

ID AAR35207 standard; protein; 3014 AA.

XX

AC AAR35207;

XX

DT 29-JUL-1993 (first entry)

XX

DE Hepatitis C virus protein.

XX

KW HCV; detection; antigen; vaccine; recombinant.

XX

OS Hepatitis C virus.

XX

FN JP05068563-A.

XX

PD 23-MAR-1993.

XX

PF 17-JUL-1991; 91JP-00203884.

XX

PR 17-JUL-1991; 91JP-00203884.

XX

PA (KAGA) KAGAKU OYOBI KESSEI RYOHU.

XX

DR WPI; 1993-130639/16.

DR

DR N-PSDB; AAQ38959.

XX

XX Nucleotide sequence encoding hepatitis C virus polypeptide - is useful

PT for detecting HCV infection, esp. in Japan, and as vaccine against HCV.

PT

XX Claim 2; Page 6-17; 17pp; Japanese.

PS

XX RNA was extracted from the plasma of Japanese patients whose HBs antigen

CC was negative and with a GPT over 100. cDNA was synthesised from the RNA

CC and cloned into lambda gt11 and screened using HCV infected chimpanzee

CC plasma to isolate HCV clones. The DNA sequence is useful in detection of

CC HCV virus. The polypeptide it produces may be used as an antigen in the

CC prep. of HCV vaccine

XX

SQ Sequence 3014 AA;

Query Match 10.2%; Score 12; DB 2; Length 3014;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

us-09-638-693a-36_copy_16_133.olg.rag

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

RESULT 436

AAR54099

ID AAR54099 standard; protein; 3014 AA.

XX

AC AAR54099;

XX

DT 09-FEB-1995 (first entry)

XX

DE NANBHV E1/E2 protein.

XX

DE E1/E2 protein; non-A, non-B hepatitis virus; NANBHV; signal peptide;

KW heterogenic; virus; transformation; insect cell; antigen; anti-NANBHV;

KW antibody; hepatitis C virus; HCV; vaccine.

XX

OS Non-A.

OS non-B hepatitis virus.

XX

EH Key Location/Qualifiers

FT Peptide 832..847

FT /note= "Peptide fragment not given in the specification,

FT but encoded by the given cDNA sequence"

FT Peptide 1296..1311

FT /note= "Peptide fragment not given in the specification,

FT but encoded by the given cDNA sequence"

FT Peptide 1760..1775

FT /note= "Peptide fragment not given in the specification,

FT but encoded by the given cDNA sequence"

FT Peptide 2688..2703

FT /note= "Peptide fragment not given in the specification,

FT but encoded by the given cDNA sequence"

XX

XX JP06141873-A.

XX

XX 24-MAY-1994.

XX

XX 13-MAR-1992; 92JP-00089371.

XX

XX 13-MAR-1992; 92JP-00089371.

XX

XX (KAGA) ZH KAGAKU & KESSEI RYOHU KENKYUSHO.

XX

XX WPI; 1994-205030/25.

DR

DR N-PSDB; AAQ64175.

XX

XX Virus vector contg hepatitis C virus and signal sequence - useful in

PT vaccines and in immunological detection.

PT

XX Disclosure; Page 7-19; 23pp; Japanese.

PS

XX This sequence represents the E1/E2 protein from non-A, non-B hepatitis

CC virus (NANBHV). The cDNA encoding this sequence may be linked to a signal

CC peptide (AAR54100) from a heterogenic virus so the the NANBHV protein may

CC be expressed by a transformed insect cell. This protein may be used as an

CC antigen in the generation of anti-NANBHV antibodies. These antibodies may

CC be used in a hepatitis C vaccine

XX

SQ Sequence 3014 AA;

Query Match 10.2%; Score 12; DB 2; Length 3014;

Best Local Similarity 100.0%; Pred. No. 0.053;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675

RESULT 437

ADX40822
ID ADX40822 standard; protein; 3019 AA.
XX
AC ADX40822;
XX
DT 21-APR-2005 (first entry)
XX
XX HCV polymerase protein #45.
DE
XX Immune stimulation; polymerase; enzyme.
KW
XX Hepatitis C virus.
OS
XX WO2005012502-A2.
PN
XX 10-FEB-2005.
PD
XX 29-MAR-2004; 2004WO-US009510.
PP
XX 28-MAR-2003; 2003US-0458026P.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
PI
XX WPI; 2005-132661/14.
XX
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
XX Disclosure; Page 388-440; 458pp; English.

The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.

XX
SQ Sequence 3019 AA;

Query Match 10.2%; Score 12; DB 9; Length 3019;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1669 GGVLAALAAAYCL 1680

RESULT 438

ADD67962
ID ADD67962 standard; protein; 3090 AA.
XX
AC ADD67962;
XX
DT 15-JAN-2004 (first entry)
XX
XX EMCV internal ribosome entry site peptide.

XX antiviral; hepatitis C virus; HCV; viral replication inhibitor;
KW replication competent HCV; 3' non-translated region; EMCV; IRES;
KW internal ribosome entry site.
XX
XX Encephalomyocarditis virus.
OS
XX US2003125541-A1.
PN
XX

PD 03-JUL-2003.

XX 27-SEP-2002; 2002US-00259275.
PF
XX

XX 23-DEC-1999; 99US-0171909P.
PR

XX 23-DEC-2000; 2000US-00747419.
PR

XX 27-SEP-2001; 2001US-0325236P.
PR

XX 13-NOV-2001; 2001US-0338123P.
XX

XX (TEXA) UNIV TEXAS SYSTEM.
PA

XX Lemon SM, Yi M;
PI

XX WPI; 2003-811006/76.
DR

XX Identifying a compound that inhibits replication of a hepatitis C virus
PT (HCV) RNA comprises contacting a cell comprising a replication competent
PT HCV RNA containing a heterologous polynucleotide encoding a
PT transactivator, with a compound.

XX Disclosure; Fig 24; 95pp; English.
PS

XX The invention describes a method of identifying a compound that inhibits
CC replication of a hepatitis C virus (HCV) RNA. The method comprises
CC contacting a cell comprising a replication competent HCV RNA containing a
CC heterologous polynucleotide having a first coding sequence encoding a
CC transactivator, with a compound. The method is useful for identifying a
CC compound that inhibits replication of HCV RNA. The kit is useful for
CC detecting replication competent HCV RNA. This sequence represents the
CC internal ribosome entry site (IRES) of Encephalomyocarditis virus (EMCV)
CC used in the creation of a modified HCV cDNA.

XX Sequence 3090 AA;
SQ

Query Match 10.2%; Score 12; DB 7; Length 3090;
Best Local Similarity 100.0%; Pred. No. 0.054;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1665 GGVLAALAAAYCL 1676

RESULT 439
AEB17115
ID AEB17115 standard; protein; 3091 AA.
XX
AC AEB17115;
XX
DT 22-SEP-2005 (first entry)
XX
XX Hepatitis C virus strain HCV-N polypeptide, SEQ ID NO: 40 #2.

XX RNA virus; microorganism; hepatitis C virus infection; antiinflammatory;
KW hepatotropic; virucide; gastrointestinal disease; infection; polypeptide.
XX
XX Hepatitis C virus; Japanese genotype 1b (HCV-N).

XX Key Location/Qualifiers
FT Region 1. .191
FT Region /note= "Core protein"
FT Region 192. .383
FT Region /note= "Envelope protein (E1)"
FT Region 384. .747
FT Region /note= "Envelope protein (E2)"
FT Region 748. .810
FT Region /note= "p7 protein"
FT Region 811. .1027
FT Region /note= "NS2 protein"
FT Region 1028. .1658
FT Region /note= "NS3 protein"
FT Region 1659. .1712
FT Region /note= "NS4A protein"

PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT
XX Disclosure; Page 175; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 15 AA;
Query Match 9.3%; Score 11; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0056;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAVC 28
DB 5 GGVLAALAAVC 15
|||||
RESULT 445
AAR63320
ID AAR63320 standard; protein; 20 AA.
AC AAR63320;
XX
XX 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 09-AUG-1995 (first entry)
XX
XX Peptide fragment of hepatitis C virus type 3.
DE
XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
XX classification; immunisation; prophylaxis; serotyping.
XX
XX Hepatitis C virus; isolate HD10.
OS
XX WO9425601-A2.
PN
XX 10-NOV-1994.
PD
XX 27-APR-1994; 94WO-EP001323.
PF
XX 27-APR-1993; 93EP-00401099.
PR
XX 05-AUG-1993; 93EP-00402019.
PR
XX (INNO-) INNOGENETICS NV SA.
PA
XX Maertens G, Stuyver L;
XX
XX WPI; 1994-358277/44.
XX
XX New polynucleotide sequences from hepatitis C virus - and related
PT vectors, polypeptide(s) and antibodies, useful for immunisation,
PT treatment, diagnosis and typing of HCV isolates.
XX
XX Disclosure; Page 167-168; 404pp; English.
PS
XX Compositions comprising at least 5, and pref. 8 or more contiguous
CC nucleotides selected from an HCV type 3 genomic sequence, more
CC particularly (i) the region spanning positions 417-957 of the Core/E1
CC

CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
CC genomic sequence, a type 4 genomic sequence; or the coding region of
CC subtype 5a, may be used as primers to amplify nucleic acid from an
CC isolate belonging to a specific genotype, or as a probe for specific
CC detection/classification of nucleic acid. Polypeptides encoded by the
CC nucleotides in such compositions may be used for immunisation against
CC HCV, for the detection of antibodies directed against HCV and for
CC serotyping. This polypeptide corresponds to positions 1724-1743 of HCV
CC type 3. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 20 AA;
Query Match 9.3%; Score 11; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 LGLLQRATQQQ 97
DB 10 LGLLQRATQQQ 20
|||||
RESULT 446
AAY67803
ID AAY67803 standard; peptide; 35 AA.
XX
XX AAY67803;
AC
XX 23-MAR-2000 (first entry)
DT
XX Peptide #203 for detecting hepatitis C virus infection.
DE
XX Hepatitis C virus; HCV; increased structural stability; NS4 region;
XX diagnostic antigen.
KW
XX Synthetic.
OS
XX WO9962945-A2.
PN
XX 09-DEC-1999.
PD
XX 04-JUN-1999; 99WO-US012446.
PF
XX 05-JUN-1998; 98US-0088229P.
PR
XX 01-SEP-1998; 98US-0098705P.
PR
XX 15-SEP-1998; 98US-0100422P.
PR
XX 28-JAN-1999; 99WO-US001726.
PR
XX (PEPT-) PEPTIDE SOLUTIONS INC.
PA
XX Chowdhury MA, Bernstein D, Mutsaers MA;
PI
XX WPI; 2000-086953/07.
XX
XX Improving properties of peptides for use as diagnostic antigens or for
PT preventing or treating infections.
PT
XX Claim 55; Page 73; 83pp; English.
PS
XX This is a peptide related to the immunoreactive region of the NS4 region
CC of hepatitis C virus (HCV). The peptide is useful for detecting HCV
CC infection. The invention relates to peptides derived from HCV and also
CC HIV-1 which have been modified for use as diagnostic antigens in the
CC treatment or prevention of infection. The structural stability of the
CC peptides can be increased in four different ways; through the replacement
CC of a hydrophobic amino acid with a less hydrophobic amino acid; through
CC an increase in the amount of secondary structure (i.e. alpha helix) in
CC the peptide; through the removal of a positive charge from the peptide,
CC or through the constraint of the epitopic sequence via the formation of a
CC

covalent crosslink. Modified peptides of the invention are used to detect infectious agents specifically HCV. Other detectable agents include HIV-1 Group O viruses; human T-cell lymphotropic virus-I or -II; and the causative agent of syphilis. The peptides can be used for prevention or treatment of infections (e.g. as vaccines, or where expressed from a transgene). More generally almost any peptide can be similarly modified, e.g. cytokines or interferons; major histocompatibility complex antigens; hormones; growth factors; tumour markers or suppressors, or antigens from many other pathogens

CC Sequence 35 AA;

Query Match 9.3%; Score 11; DB 3; Length 35;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEVLYQQY 59

Db 24 VPDKEVLYQQY 34

RESULT 447

AAE20477

ID AAE20477 standard; protein; 3010 AA.

XX

AC AAE20477;

DT 01-JUL-2002 (first entry)

XX

XX HCV-S1 full-length polyprotein.

XX

Nucleic acid construct; expression cassette; non-coding region; NCR;

untranslated region; UTR; anti-viral drug; drug resistance; HCV-S1;

Hepatitis C virus.

XX

OS Hepatitis C virus.

XX

FN WO200208447-A2.

XX

PD 31-JAN-2002.

XX

XX 20-JUL-2001; 2001WO-IL000669.

XX

PR 24-JUL-2000; 2000US-022048P.

XX

XX (MOLE-) INST MOLECULAR & CELL BIOLOGY.

PA (EHLR/) EHLRICH G.

XX

PI Tan YH, Lim SP, Lim SG, Hong WJ;

XX

DR WPI; 2002-280605/32.

XX

DR N-PSDB; AAD33038.

XX

Novel nucleic acid construct useful for detecting the presence of RNA virus, comprises an expression cassette and a promoter operably linked to expression cassette for minus strand RNA transcription of the cassette.

XX Example 1; Page 70-81; 81pp; English.

XX

The invention relates to nucleic acid construct which comprises an expression cassette including a first polynucleotide region including a 5' non-coding region (NCR) sequence of an RNA virus and at least an N-terminal portion of a coding sequence of RNA virus, a second polynucleotide region including a 3' untranslated region (UTR) sequence of the RNA virus and at least a C-terminal portion of a coding sequence of the RNA virus and a third polynucleotide region encoding a reporter molecule, flanked by first and second polynucleotide regions; and a promoter sequence being operatively linked to expression cassette in a manner so as to enable a transcription of a minus strand RNA molecule from the expression cassette. Nucleic acid construct of the invention is useful for detecting the presence of an RNA virus in a cell. It is also useful for screening anti-viral drugs and determining drug resistance of an RNA virus. The present sequence is Hepatitis C virus (HCV) isolate HCV

CC -S1 full-length polyprotein

XX

SQ Sequence 3010 AA;

Query Match 9.3%; Score 11; DB 5; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALAAAYCL 29

Db 1665 GVLAALAAAYCL 1675

XX

RESULT 448

ADX40809

ID ADX40809 standard; protein; 3022 AA.

XX

AC ADX40809;

XX

DT 21-APR-2005 (first entry)

XX

DE HCV polymerase protein #32.

XX

KW Immune stimulation; polymerase; enzyme.

XX

OS Hepatitis C virus.

XX

FN WO2005012502-A2.

XX

PD 10-FEB-2005.

XX

XX 29-MAR-2004; 2004WO-US009510.

XX

PR 28-MAR-2003; 2003US-0458026P.

XX

XX (EPIM-) EPIMUNE INC.

XX

PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;

XX

XX WPI; 2005-132661/14.

XX

PT Identifying a candidate peptide epitope, which induces a HLA class I CTL

response comprises identifying variants of a peptide epitope 8-11 amino

acids in length comprising primary anchor residues of the same HLA class

I binding motif.

XX

PS Disclosure; Page 388-440; 458pp; English.

XX

CC The invention relates to a method of identifying a candidate peptide

epitope which induces an HLA class I CTL response against variants of the

peptide epitope, comprising identifying, from a particular antigen of an

infectious agent, variants of a peptide epitope comprising primary anchor

residues of the same HLA class I binding motif. The method is useful for

identifying a candidate peptide epitope, which induces an HLA class I CTL

response against variants of the peptide epitope. This sequence

represents an HCV polymerase protein used in the scope of the invention.

XX

SQ Sequence 3022 AA;

Query Match 9.3%; Score 11; DB 9; Length 3022;

Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 YQYDEMEECs 66

Db 1706 YQYDEMEECs 1716

XX

RESULT 449

AAR73113

ID AAR73113 standard; peptide; 10 AA.

XX

AC AAR73113;

```

XX 25-MAR-2003 (revised)
DT 16-JUN-1995 (first entry)
XX
XX Antigen fragment 10 from HCV has binding affinity for HLA-A2.1.
XX
XX antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
KW plasma specific antigen; hepatitis B virus; Epstein Barr;
KW human immunodeficiency virus; human papilloma virus; p53; c-ERB2; MAGE-1;
KW melanoma antigen-1; core antigen; surface antigen;
KW pharmaceutical composition; in vivo; ex vivo; therapeutic; diagnostic;
KW MHC class I molecule; major histocompatibility complex; HLA-A2.1; 9mer;
KW 10mer; anchor; human leukocyte antigen; PLP; 8mer; algorithm prediction;
KW MBP; CMV; cytomegalovirus; HSV; herpes simplex virus; influenza A; M1.
XX
XX Hepatitis C virus.
XX
XX WO9420127-A1.
XX
XX 15-SEP-1994.
XX
XX 04-MAR-1994; 94WO-US002353.
XX
XX 05-MAR-1993; 93US-00027146.
XX
XX 04-JUN-1993; 93US-00073205.
XX
XX 29-NOV-1993; 93US-00159184.
XX
XX (CYTE-) CYTEL CORP.
XX
XX Grey HM, Sette A, Sidney J, Kast W;
XX
XX WPI; 1994-302678/37.
XX
XX Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
XX treatment or prophylaxis of cancer, virus infection or autoimmune
XX diseases.
XX
XX Disclosure; Page 90; 138pp; English.
XX
XX AAR73058-121 are potential peptide binders of HLA-A2.1 motif. Using
XX motifs disclosed in the invention, these peptides were screened for
XX further motifs. Only peptides with binding affinity of at least 1%
XX (binding affinity is expressed as an IC50 value) as compared to the
XX standard peptide (AAR71293) in assays. This peptide from hepatitis C
XX virus has an binding value of 0.010. The peptides of the invention can
XX induce cytotoxic T lymphocytes which can react with target cells. They
XX can be used for the treatment or prophylaxis of cancer, eg. prostate
XX cancer or lymphoma, etc. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 10 AA;
XX
XX Query Match 8.5%; Score 10; DB 2; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.035;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 20 VLAAALAYCL 29
XX |||||
XX 1 VLAAALAYCL 10
XX
XX RESULT 450
XX AAR61533
XX ID AAR61533 standard; peptide; 10 AA.
XX
XX AC AAR61533;
XX
XX 25-MAR-2003 (revised)
DT 11-MAY-1995 (first entry)
XX
XX Peptide fragment (1.0889) of HCV binds HLA-A2.1.
XX
XX antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
KW plasma specific antigen; hepatitis B virus; Epstein Barr;

```

```

KW human immunodeficiency virus; human papilloma virus; p53; c-ERB2; MAGE-1;
KW melanoma antigen-1; core antigen; surface antigen;
KW pharmaceutical composition; in vivo; ex vivo; therapeutic; diagnostic;
KW MHC class I molecule; major histocompatibility complex; HLA-A2.1; 9mer;
KW 10mer; anchor; human leukocyte antigen.
XX
XX Hepatitis C virus.
XX
XX WO9420127-A1.
XX
XX 15-SEP-1994.
XX
XX 04-MAR-1994; 94WO-US002353.
XX
XX 05-MAR-1993; 93US-00027146.
XX
XX 04-JUN-1993; 93US-00073205.
XX
XX 29-NOV-1993; 93US-00159184.
XX
XX (CYTE-) CYTEL CORP.
XX
XX Grey HM, Sette A, Sidney J, Kast W;
XX
XX WPI; 1994-302678/37.
XX
XX Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
XX treatment or prophylaxis of cancer, virus infection or autoimmune
XX diseases.
XX
XX Example 5; Page 108; 138pp; English.
XX
XX AAR59496-R61666 are immunogenic 10mer peptides that contain a HLA-A2.1
XX binding motif. These peptides bind HLA-A2.1 and have a binding affinity
XX of at least 1% as compared to a reference peptide (AAR71293). AAR61533
XX has an IC50 of 0.011 and the sequence occurs at position 1666 in the HCV
XX LORF protein. The peptides of the invention can induce cytotoxic T
XX lymphocytes which can react with target cells. They can be used for the
XX treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 10 AA;
XX
XX Query Match 8.5%; Score 10; DB 2; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.035;
XX Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 20 VLAAALAYCL 29
XX |||||
XX 1 VLAAALAYCL 10
XX
XX RESULT 451
XX AAR78955
XX ID AAR78955 standard; peptide; 10 AA.
XX
XX AC AAR78955;
XX
XX 25-MAR-2003 (revised)
DT 01-APR-1996 (first entry)
XX
XX HCV NS4 1666-1675 cytotoxic T lymphocyte epitope.
XX
XX HCV NS4 1666-1675; cytotoxic T; CTL; epitope; helper T; HTL; cell;
KW lymphocyte; antigens; treatment; disease prevention; hepatitis C; non-A;
KW non-B.
XX
XX Hepatitis C virus.
XX
XX WO9522317-A1.
XX
XX 24-AUG-1995.
XX
XX 16-FEB-1995; 95WO-US002121.
XX

```

PR 16-FEB-1994; 94US-00197484.
XX (CYTE-) CYTEL CORP.
XX Vitiello MA, Chesnut RW, Sette AD, Celis E, Grey H;
XX WPI; 1995-302545/39.
XX Compen. inducing cytotoxic T lymphocyte response to pref. viral,
XX bacterial, parasitic or tumour antigens - useful in the treatment and
XX prevention of diseases associated with the antigen e.g. hepatitis B.
PS Example 12; Page 70; 109pp; English.
XX
CC A compen. which induces a cytotoxic T lymphocyte (CTL) response to a
CC hepatitis C virus (HCV) antigen (Ag) in a mammal comprises, a HCV CTL Ag
CC response inducing peptide (i.e. AAR78941-R78955) and a lipid conjugated
CC helper T cell inducing peptide. The compen. is useful in the treatment
CC and prevention of hepatitis C. (Updated on 25-MAR-2003 to correct PI
CC field.)
XX
XX Sequence 10 AA;
SQ

Query Match 8.5%; Score 10; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAAALAAAYCL 29
DB 1 VLAAALAAAYCL 10

RESULT 452
AAJ02685
ID AAJ02685 standard; peptide; 10 AA.
XX
AC AAJ02685;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #2676.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
OS Hepatitis C virus.
XX
XX WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 127; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 10 AA;
XX

Query Match 8.5%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAAALAAAYCL 29
DB 1 VLAAALAAAYCL 10

RESULT 454
AAJ00646
ID AAJ00646 standard; peptide; 10 AA.
XX
AC AAJ00646;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #637.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX

Query Match 8.5%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAAALAAAYCL 29
DB 1 VLAAALAAAYCL 10

RESULT 454
AAJ00646
ID AAJ00646 standard; peptide; 10 AA.
XX
AC AAJ00646;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #637.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX

```
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX XX
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 116; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 10 AA;

Query Match 8.5%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAY 27
Db 1 GGVLAALAAAY 10

RESULT 456
AAJ01923
ID AAJ01923 standard; peptide; 10 AA.
XX AC AAJ01923;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1914.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX XX WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 148; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 10 AA;

Query Match 8.5%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.035;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAALAAAYCL 29
Db 1 VLAAALAAAYCL 10

RESULT 457
AAJ02192
ID AAJ02192 standard; peptide; 10 AA.
XX AC AAJ02192;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #2183.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX XX 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 155; 214pp; English.
XX XX
```


CC binding selected major histocompatibility complex (MHC) molecules and
 CC inducing an immune response. The pharmaceutical compositions are useful
 CC for the treatment of human viral diseases (e.g. prostate cancer,
 CC hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical carcinoma,
 CC lymphoma, cytomegalovirus and condyloma acuminatum), cancers or
 CC autoimmune diseases, and to relieve the symptoms of, treat or prevent the
 CC occurrence or reoccurrence of autoimmune diseases e.g. multiple sclerosis
 CC (MS), rheumatoid arthritis, Sjogren syndrome, scleroderma, polymyositis,
 CC dermatomyositis, systemic lupus erythematosus, juvenile rheumatoid
 CC arthritis, ankylosing spondylitis, myasthenia gravis, bullous pemphigoid,
 CC pemphigus, glomerulonephritis, Goodpasture's syndrome, autoimmune
 CC haemolytic anaemia, Hashimoto's disease, pernicious anaemia, idiopathic
 CC thrombocytopenic purpura, Grave's disease, and Addison's disease. The
 CC compositions are capable of specifically binding glycoproteins encoded by
 CC HLA-A2.1 allele and inducing T cell activation in T cells restricted by
 CC the A2.1 allele. This is the amino acid sequence of a HLA-A2.1 binding
 CC peptide.

XX Sequence 10 AA;

Query Match 8.5%; Score 10; DB 8; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.035;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAALAYCL 29
 Db 1 VLAAALAYCL 10
 |||||

RESULT 459

AAJ02897
 ID AAJ02897 standard; peptide; 11 AA.

XX AC AAJ02897;
 XX DT 02-JUL-2001 (first entry)
 XX DE Hepatitis C virus epitope #2888.
 XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 XX KW antiviral.

OS Hepatitis C virus.

XX WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccine against hepatitis C virus.

XX PS Disclosure; Page 171; 214pp; English.

XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX Sequence 11 AA;

Query Match 8.5%; Score 10; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.038;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAY 27
 Db 2 GGVLAAALAY 11
 |||||

RESULT 460

AAJ02554
 ID AAJ02554 standard; peptide; 11 AA.

XX AC AAJ02554;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #2545.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 XX KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccine against hepatitis C virus.

XX PS Disclosure; Page 163; 214pp; English.

XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX Sequence 11 AA;

Query Match 8.5%; Score 10; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.038;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAY 27
 Db 2 GGVLAAALAY 11
 |||||

RESULT 461

AAW85432
 ID AAW85432 standard; peptide; 15 AA.

XX AC AAW85432;

XX DT 16-FEB-1999 (first entry)

XX DE Helper T-cell class II peptide derived from NS4 protein.

XX KW Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1; DR7;
 XX KW cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
 XX KW acquired immune deficiency syndrome; malaria; cancer;
 XX KW allograft rejection; allergy; Lyme disease; hepatitis;
 XX KW post-streptococcal endocarditis; glomerulonephritis;

```

KW food hypersensitivity.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
PN WO9832456-A1.
XX
PD 30-JUL-1998.
XX
XX
XX 23-JAN-1998; 98WO-US001373.
XX
XX 23-JAN-1997; 97US-0036713P.
PR 07-FEB-1997; 97US-0037432P.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Sette A, Sidney J, Southwood S;
XX
XX WPI; 1998-427679/36.
XX
XX Composition containing peptide that induces cytotoxic T lymphocyte
PT response, and helper peptide - can bind to human leucocyte antigen
PT alleles, used to treat or prevent cancers, parasitic infections and
PT autoimmune disease.
XX
XX Disclosure; Page 42; 51pp; English.
XX
XX AA085284-451 represent helper T-cell class II peptides, which can bind to
CC the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides are
CC used in the course of the invention. The specification describes peptides
CC that that induce a cytotoxic T lymphocyte (CTL) response, and T-helper
CC peptides, that are used together to generate a CTL response for the
CC treatment or prevention of viral, fungal, bacterial or parasitic
CC infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC cancer or condyloma acuminatum). Helper T-cell peptides may be used alone
CC to induce a helper T cell response, e.g. in cases of autoimmune disease,
CC allograft rejection, allergy, Lyme disease, hepatitis, post-streptococcal
CC endocarditis, glomerulonephritis and food hypersensitivity
XX
XX Sequence 15 AA;
SQ
Query Match 8.5%; Score 10; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAY 27
Db |||||
6 GGVLAALAAAY 15
RESULT 462
AAJ03166
ID AAJ03166 standard; peptide; 15 AA.
XX
XX AAJ03166;
AC
XX
XX 02-JUL-2001 (first entry)
DT
XX
XX Hepatitis C virus epitope #3157.
DE
XX
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
XX Hepatitis C virus.
OS
XX
XX WO200121189-A1.
PN
XX
XX 29-MAR-2001.
PD
XX
XX 19-JUL-2000; 2000WO-US019774.
PF
XX
XX 19-JUL-1999; 99US-00357737.
PR
XX
XX (EPIM-) EPIMMUNE INC.
PA
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
XX WPI; 2001-308046/32.
DR
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT
XX
XX Example 5; Page 198; 214pp; English.
PS
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 15 AA;
SQ
Query Match 8.5%; Score 10; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAY 27
Db |||||
6 GGVLAALAAAY 15
RESULT 463
AAJ04008
ID AAJ04008 standard; peptide; 15 AA.
XX
XX AAJ04008;
AC
XX
XX 02-JUL-2001 (first entry)
DT
XX
XX Hepatitis C virus epitope #3999.
DE
XX
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
XX Hepatitis C virus.
OS
XX
XX WO200121189-A1.
PN
XX
XX 29-MAR-2001.
PD
XX
XX 19-JUL-2000; 2000WO-US019774.
PF
XX
XX 19-JUL-1999; 99US-00357737.
PR
XX
XX (EPIM-) EPIMMUNE INC.
PA
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
XX WPI; 2001-308046/32.
DR
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT
XX
XX Example 5; Page 198; 214pp; English.
PS
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 15 AA;
SQ
Query Match 8.5%; Score 10; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.049;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAY 27
Db |||||
6 GGVLAALAAAY 15
RESULT 464
AAJ04008
ID AAJ04008 standard; peptide; 15 AA.
XX
XX AAJ04008;
AC
XX
XX 02-JUL-2001 (first entry)
DT
XX
XX Hepatitis C virus epitope #3999.
DE
XX
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
XX Hepatitis C virus.
OS
XX
XX WO200121189-A1.
PN
XX
XX 29-MAR-2001.
PD
XX
XX 19-JUL-2000; 2000WO-US019774.
PF
XX
XX 19-JUL-1999; 99US-00357737.
PR
```

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAY 27
 |||||
 Db 6 GGVLAALAAAY 15

RESULT 464
 AAJ03473
 ID AAJ03473 standard; peptide; 15 AA.
 XX AC AAJ03473;
 XX DT 02-JUL-2001 (first entry)
 XX DE Hepatitis C virus epitope #3464.
 XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 XX KW antiviral.
 XX OS Hepatitis C virus.
 XX PN WO200121189-A1.
 XX PD 29-MAR-2001.
 XX PF 19-JUL-2000; 2000WO-US019774.
 XX PR 19-JUL-1999; 99US-00357737.
 XX PA (EPIM-) EPIMMUNE INC.
 XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 XX PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX DR WPI; 2001-308046/32.
 XX PT A new composition useful as a vaccines against hepatitis C virus.
 XX PF Disclosure; Page 178; 214pp; English.
 XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX SQ Sequence 15 AA;

Query Match 8.5%; Score 10; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAY 27
 |||||
 Db 6 GGVLAALAAAY 15

RESULT 465
 ADM36129
 ID ADM36129 standard; peptide; 15 AA.
 XX AC ADM36129;
 XX DT 10-MAR-2005 (first entry)
 XX DE HLA binding epitope #6879.
 XX KW Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
 XX KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
 XX KW viral disease; cancer.
 XX OS Unidentified.

XX PN WO2003040165-A2.
 XX PD 15-MAY-2003.
 XX PF 18-OCT-2001; 2001WO-US051650.
 XX PR 19-OCT-2000; 2000US-0242350P.
 XX PR 20-APR-2001; 2001US-0285624P.
 XX PA (EPIM-) EPIMMUNE INC.
 XX PI Sette A, Sidney J, Southwood S;
 XX DR WPI; 2003-441519/41.
 XX PT New composition comprising at least one peptide having allele-specific
 XX PT binding motifs for HLA, useful for preventing, treating or diagnosing
 XX PT viral diseases and cancer.
 XX PF Claim 1; Page 52-379; 382pp; English.
 XX CC The invention relates to a composition comprising at least one peptide
 CC having an isolated, prepared epitope selected from any of the sequences
 CC from 30 lists given in the specification. Also disclosed is a method for
 CC inducing a cytotoxic T cell response against a pre-selected antigen in a
 CC patient expressing a specific MHC class I allele by contacting cytotoxic
 CC T cells from the patient with the composition cited above. The
 CC composition comprises an epitope that is joined by an amino acid linker.
 CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
 CC bound to an HLA molecule on the antigen-presenting cell, where when an A2
 CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
 CC binds to a complex of the HLA molecule and the epitope. Specifically
 CC claimed are peptides having allele-specific binding motifs for HLA. The
 CC compositions and methods are useful for preventing, treating or
 CC diagnosing viral diseases and cancer. The peptide epitopes are useful as
 CC diagnostic agents for evaluating immune responses, for making antibodies
 CC and for evaluating efficacy of a vaccine. Sequences given in ADM29251-
 CC ADM37745 represent epitopes of the invention as given in Tables 2-31.
 XX SQ Sequence 15 AA;

Query Match 8.5%; Score 10; DB 7; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAY 27
 |||||
 Db 6 GGVLAALAAAY 15

RESULT 466
 ADM35960
 ID ADM35960 standard; peptide; 15 AA.
 XX AC ADM35960;
 XX DT 10-MAR-2005 (first entry)
 XX DE HLA binding epitope #6710.
 XX KW Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
 XX KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
 XX KW viral disease; cancer.
 XX OS Unidentified.
 XX PN WO2003040165-A2.
 XX PD 15-MAY-2003.
 XX PF 18-OCT-2001; 2001WO-US051650.
 XX SQ

PR 19-OCT-2000; 2000US-0242350P.
 PR 20-APR-2001; 2001US-0285624P.
 PA (EPIM-) EPIMUNE INC.
 XX Sette A, Sidney J, Southwood S;
 PI WPI; 2003-441519/41.
 DR
 XX
 PT New composition comprising at least one peptide having allele-specific
 PT binding motifs for HLA, useful for preventing, treating or diagnosing
 PT viral diseases and cancer.
 XX
 PS Claim 1; Page 52-379; 382pp; English.
 XX
 CC The invention relates to a composition comprising at least one peptide
 CC having an isolated, prepared epitope selected from any of the sequences
 CC from 30 lists given in the specification. Also disclosed is a method for
 CC inducing a cytotoxic T cell response against a pre-selected antigen in a
 CC patient expressing a specific MHC class I allele by contacting cytotoxic
 CC T cells from the patient with the composition cited above. The
 CC composition comprises an epitope that is joined by an amino acid linker.
 CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
 CC bound to an HLA molecule on the antigen-presenting cell, where when an A2
 CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
 CC binds to a complex of the HLA molecule and the epitope. Specifically
 CC claimed are peptides having allele-specific binding motifs for HLA. The
 CC compositions and methods are useful for preventing, treating or
 CC diagnosing viral diseases and cancer. The peptide epitopes are useful as
 CC diagnostic agents for evaluating immune responses, for making antibodies
 CC and for evaluating efficacy of a vaccine. Sequences given in ADW29251-
 CC ADM37745 represent epitopes of the invention as given in Tables 2-31.
 XX
 SQ Sequence 15 AA;
 Query Match 8.5%; Score 10; DB 7; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAY 27
 DB 6 GGVLAALAAAY 15
 RESULT 467
 ADV23001
 ID ADV23001 standard; peptide; 15 AA.
 XX
 AC ADV23001;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE HCV H77 immunogenic peptide #242.
 XX
 KW Vaccine; virucide; antigen; autoimmune disease; infection;
 KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
 KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
 KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
 KW hodgkin's lymphoma.
 XX
 OS Hepatitis C virus strain H77.
 XX
 FN WO2004108753-A1.
 XX
 PD 16-DEC-2004.
 XX
 XX 10-JUN-2004; 2004WO-AU000775.
 PF
 XX 10-JUN-2003; 2003AU-00902875.
 XX
 PR 25-MAR-2004; 2004AU-00901589.
 PR
 XX (UTME) UNIV MELBOURNE.
 PA
 XX

Kent SJ;
 WPI; 2005-031657/03.
 Use of at least one set of peptides in the preparation of a medicament
 for modulating an immune response, and for treating cancer or yeast,
 viral, bacterial, protozoal and mycoplasma infections.
 Disclosure; SEQ ID NO 1421; 645pp; English.
 The invention relates to the use of at least one set of peptides in the
 preparation of a medicament for modulating an immune response, where
 individual peptides of a respective set comprise different portions of an
 amino acid sequence corresponding to a single polypeptide of interest and
 display partial sequence identity or similarity to at least one other
 peptide of the same set of peptides (i.e. they are overlapping). Also
 included are an antigen-presenting cell which has been contacted with the
 peptides above and thus presents the peptides, a population of such
 antigen-presenting cells, a process for producing antigen-presenting
 cells for modulating an immune response to a polypeptide of interest, a
 method for producing antigen-specific lymphocytes, a composition
 comprising at least one set of the peptides (and a carrier and/or
 diluent), a method for modulating an immune response to a polypeptide of
 interest comprising administering to a patient in need at least one set
 of the peptides, a method for treatment and/or prophylaxis of a disease
 or condition associated with the presence of a polypeptide of interest
 and a composition of matter for modulating an immune response in a
 subject to a target antigen. The polypeptide of interest is also a
 disease- or condition-associated polypeptide that is a polypeptide
 produced by a pathogenic organism or a cancer, and produced by a
 pathogenic organism selected from yeast, viruses, bacteria, helminths,
 protozoans and mycoplasmas. The disease- or condition-associated
 polypeptide is produced by a cancer selected from melanoma, lung cancer,
 breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
 cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
 lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
 antigen-presenting cells or their precursors are useful in the
 preparation of a medicament for the treatment of a disease or condition
 in a subject, which disease or condition is associated with the presence
 of aberrant expression of a target antigen, where the antigen-presenting
 cells or their precursors have not been subjected to activating
 conditions but have been contacted with an antigen that corresponds to
 the target antigen to express a processed or modified form of the antigen
 for presentation to the subject's immune system. The present sequence is
 one of a set of overlapping immunogenic peptides derived from a Hepatitis
 C virus protein.

Query Match 8.5%; Score 10; DB 9; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.049;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAAAYCLS 30
 DB 1 LAALAAAYCLS 10
 RESULT 468
 AAR20713
 ID AAR20713 standard; protein; 303 AA.
 XX
 AC AAR20713;
 XX
 DT 27-AUG-2003 (revised)
 DT 06-MAY-1992 (first entry)
 XX
 XX C10-18 NANBH-specific antigen polypeptide.
 DE
 XX Non-A non-B hepatitis virus; recombinant; detection.
 KW
 XX Non-A.
 OS non-B hepatitis virus.
 OS

XX PN EP468657-A.
XX PD 29-JAN-1992.
XX PF 09-JUL-1990; 90JP-00180889.
XX PR 09-JUL-1990; 90JP-00180889.
XX PR 30-NOV-1990; 90JP-00339589.
XX PR 20-DEC-1990; 90JP-00413844.
XX PA (TOFU) TONEN CORP.
XX PI Maki N, Yamaguchi K, Toyoshima A, Kohara M;
XX WPI; 1992-034390/05.
XX DR N-PSDB; AAQ20624.
XX PR Non-A, non-B hepatitis-specific antigen polypeptide - for detection of
PT hepatitis virus gene or antibody directed against the virus.
XX PS Disclosure; Fig 8; 78pp; English.
XX CC The amino acid sequence is that of a non-A non-B (NANB) hepatitis
CC specific antigen polypeptide, which may be recombinantly produced. It is
CC encoded by a fragment of DNA clone Cl0-18. It can be used to prepare
CC antibodies which can be used to detect NANB hepatitis with extremely high
CC accuracy. It can also be used to detect anti-NANB hepatitis antibodies.
CC See also AAR20609-R20615 and AAR20714-R20723. (Updated on 27-AUG-2003 to
CC correct OS field.)
XX SQ Sequence 303 AA;
Query Match 8.5%; Score 10; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAY 27
Db 264 GGVLAALAAAY 273
RESULT 469
AA46886
ID AA46886 standard; peptide; 9 AA.
XX AC AAY46886;
XX DT 01-DEC-1999 (first entry)
XX DE Immunogenic peptide having a human leukocyte antigen binding motif #1497.
XX DE Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;
XX KW immune response; T cell activation; major histocompatibility complex;
XX KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;
XX KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;
XX KW vaccine; immunisation.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9945954-A1.
XX PD 16-SEP-1999.
XX PF 13-MAR-1998; 98WO-US005039.
XX PR 13-MAR-1998; 98WO-US005039.
XX PA (EPTM-) EPIMMUNE INC.
XX PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;
XX

DR WPI; 1999-551214/46.
XX New immunogenic peptides with HLA binding motif, useful in treatment and
PT diagnosis of cancers and viral diseases.
XX PS Claim 1; Page 87; 150pp; English.
XX CC AA45390 to AA48214 represent specifically claimed immunogenic peptides
CC having a human major histocompatibility complex (MHC) Class I (also known
CC as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides
CC can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2
CC or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against
CC the antigen from which the peptide is derived. Cytotoxic T lymphocytes
CC (CTLs) which destroy antigen-bearing cells are normally induced by an
CC antigen in the form of a peptide fragment bound to a HLA molecule, rather
CC than the intact foreign antigen itself, and are particularly important in
CC tumour rejection and in fighting viral infections. The peptides are
CC therefore useful therapeutically to treat or prevent viral infections and
CC cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B
CC and C, AIDS, and renal carcinoma. They can be administered as vaccines to
CC elicit an immune response in individuals susceptible or otherwise at risk
CC of viral infection or cancer, or used to treat chronic or acute
CC conditions. They are also useful diagnostically, and can be used to
CC induce a cytotoxic T cell response, by contacting a cytotoxic T cell with
CC the peptide e.g. to produce CTLs ex vivo for infusion back into a
CC patient. The polynucleotides encoding the immunogenic peptides are also
CC useful therapeutically and for immunisation as above
XX SQ Sequence 9 AA;
Query Match 7.6%; Score 9; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 GVLAAALAAAY 27
Db 1 GVLAAALAAAY 9
RESULT 470
AAJ00391
ID AAJ00391 standard; peptide; 9 AA.
XX AC AAJ00391;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #382.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPTM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccine against hepatitis C virus.
XX Disclosure; Page 110; 214pp; English.
XX

CC The present invention describes a composition comprising a prepared
 CC Hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX Sequence 9 AA;
 SQ

Query Match 7.6%; Score 9; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06; 0; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAAYCL 29
 DB 1 LAALAAYCL 9

RESULT 471

AAJ00105
 ID AAJ00105 standard; peptide; 9 AA.

XX AC
 XX AAJ00105;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #96.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Cheesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccines against hepatitis C virus.

XX PS Disclosure; Page 103; 214pp; English.

XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX Sequence 9 AA;

Query Match 7.6%; Score 9; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALAAAY 27

DB 1 GVLAALAAAY 9

RESULT 472

AAJ00900
 ID AAJ00900 standard; peptide; 9 AA.

XX AC
 XX AAJ00900;

XX

DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #891.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Cheesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccines against hepatitis C virus.

XX PS Disclosure; Page 123; 214pp; English.

XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX Sequence 9 AA;

Query Match 7.6%; Score 9; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALAAAY 27

DB 1 GVLAALAAAY 9

RESULT 473

AAJ02704
 ID AAJ02704 standard; peptide; 9 AA.

XX AC
 XX AAJ02704;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #2695.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 KW antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Cheesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;

```
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
XX
XX Disclosure; Page 166; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 9 AA;
XX
XX Query Match 7.6%; Score 9; DB 4; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 19 GVLAALAAAY 27
XX |||||
XX 1 GVLAALAAAY 9
XX
XX RESULT 474
XX AAJ01418
XX ID AAJ01418 standard; peptide; 9 AA.
XX
XX AC AAJ01418;
XX
XX DT 02-JUL-2001 (first entry)
XX
XX DE Hepatitis C virus epitope #1409.
XX
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX OS Hepatitis C virus.
XX
XX PN WO200121189-A1.
XX
XX PD 29-MAR-2001.
XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX
XX PR 19-JUL-1999; 99US-00357737.
XX
XX PA (EPIM-) EPIMMUNE INC.
XX
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX
XX DR WPI; 2001-308046/32.
XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX
XX PS Disclosure; Page 176; 214pp; English.
XX
XX CC The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX
XX SQ Sequence 9 AA;
XX
XX Query Match 7.6%; Score 9; DB 4; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 20 VLAALAAAYC 28
XX |||||
XX 1 VLAALAAAYC 9
XX
XX RESULT 476
XX AAJ02222
XX ID AAJ02222 standard; peptide; 9 AA.
XX
XX AC AAJ02222;
XX
XX DT 02-JUL-2001 (first entry)
XX
XX DE Hepatitis C virus epitope #2213.
XX
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX OS Hepatitis C virus.
XX
XX PN WO200121189-A1.
XX
XX PD 29-MAR-2001.
XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX
```

XX
PR 19-JUL-1999; 99US-00357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT Disclosure; Page 155; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
SQ

Query Match 7.6%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALAAAY 27
Db 1 GVLAALAAAY 9

RESULT 477
AAJ03317
ID AAJ03317 standard; peptide; 9 AA.
XX AAJ03317;
AC
XX 02-JUL-2001 (first entry)
DT
XX Hepatitis C virus epitope #3308.
DE
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS
XX WO200121189-A1.
PN
XX 29-MAR-2001.
PD
XX 19-JUL-2000; 2000WO-US019774.
PF
XX 19-JUL-1999; 99US-00357737.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT Disclosure; Page 177; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
SQ

Query Match 7.6%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALAAAY 27
Db 1 GVLAALAAAY 9

RESULT 477
AAJ03317
ID AAJ03317 standard; peptide; 9 AA.
XX AAJ03317;
AC
XX 02-JUL-2001 (first entry)
DT
XX Hepatitis C virus epitope #3308.
DE
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS
XX WO200121189-A1.
PN
XX 29-MAR-2001.
PD
XX 19-JUL-2000; 2000WO-US019774.
PF
XX 19-JUL-1999; 99US-00357737.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT Disclosure; Page 177; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
SQ

Query Match 7.6%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAAYCL 29
Db 1 LAALAAYCL 9

RESULT 478
AAJ03453
ID AAJ03453 standard; peptide; 9 AA.
XX AAJ03453;
AC
XX 02-JUL-2001 (first entry)
DT
XX Hepatitis C virus epitope #3444.
DE
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS
XX WO200121189-A1.
PN
XX 29-MAR-2001.
PD
XX 19-JUL-2000; 2000WO-US019774.
PF
XX 19-JUL-1999; 99US-00357737.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT Disclosure; Page 178; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
SQ

Query Match 7.6%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAAYC 28
Db 1 VLAALAAYC 9

RESULT 479
AAJ02191
ID AAJ02191 standard; peptide; 9 AA.
XX AAJ02191;
AC
XX 02-JUL-2001 (first entry)
DT
XX Hepatitis C virus epitope #2182.
DE
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX

```
OS Hepatitis C virus.
XX WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 155; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX CC
XX SQ Sequence 9 AA;

Query Match 7.6%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAAAYCL 29
Db 1 LAALAAAYCL 9

RESULT 481
AAJ01767
ID AAJ01767 standard; peptide; 9 AA.
XX AC AAJ01767;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1758.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 144; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 9 AA;

Query Match 7.6%; Score 9; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAA 26
Db 1 GGVLAALAA 9

RESULT 480
AAJ03012
ID AAJ03012 standard; peptide; 9 AA.
XX AC AAJ03012;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #3003.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 174; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
```

```
DT 02-JUL-2001 (first entry)
XX Hepatitis C virus epitope #2544.
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
XX Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 163; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX SQ Sequence 10 AA;
XX
XX Query Match 7.6%; Score 9; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.31;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAA 26
XX |||||
XX 2 GGVLAALAA 10
XX
XX RESULT 483
XX AAJ02352
XX ID AAJ02352 standard; peptide; 11 AA.
XX AC AAJ02352;
XX
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #2343.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX OS Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 163; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX SQ Sequence 10 AA;
XX
XX Query Match 7.6%; Score 9; DB 4; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 0.31;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAA 26
XX |||||
XX 2 GGVLAALAA 10
XX
XX RESULT 483
XX AAJ02352
XX ID AAJ02352 standard; peptide; 11 AA.
XX AC AAJ02352;
XX
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #2343.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX OS Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 163; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX SQ Sequence 11 AA;
XX
XX Query Match 7.6%; Score 9; DB 4; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.33;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAA 26
XX |||||
XX 3 GGVLAALAA 11
XX
XX RESULT 484
XX AAJ00438
XX ID AAJ00438 standard; peptide; 11 AA.
XX AC AAJ00438;
XX
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #429.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX OS Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 111; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX SQ Sequence 11 AA;
XX
XX Query Match 7.6%; Score 9; DB 4; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.33;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAALAA 26
XX |||||
XX 3 GGVLAALAA 11
XX
```


CC	present sequence is an epitope used in the disclosure of the invention
XX	
SQ	Sequence 15 AA;
Query Match	7.6%; Score 9; DB 4; Length 15;
Best Local Similarity	100.0%; Pred. No. 0.43;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	18 GGVLAAALAA 26
Db	7 GGVLAAALAA 15
RESULT 488	
AAJ03705	
ID	AAJ03705 standard; peptide; 15 AA.
XX	
AC	AAJ03705;
XX	
DT	02-JUL-2001 (first entry)
XX	
DE	Hepatitis C virus epitope #3696.
XX	
KW	Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif; antiviral.
KW	
OS	Hepatitis C virus.
XX	
PN	WO200121189-A1.
XX	
PD	29-MAR-2001.
XX	
Pf	19-JUL-2000; 2000WO-US019774.
XX	
PR	19-JUL-1999; 99US-00357737.
XX	
PA	(EPIM-) EPIMUNE INC.
XX	
PI	Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI	Baker DM, Celis E, Kubo RT, Grey HM;
XX	
DR	WPI; 2001-308046/32.
XX	
PT	A new composition useful as a vaccines against hepatitis C virus.
XX	
PS	Disclosure; Page 181; 214pp; English.
XX	
CC	The present invention describes a composition comprising a prepared Hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121. These are derived from HCV HLA-binding motifs. They are useful in CC vaccines for the prevention and treatment of HCV infection in humans. The present sequence is an epitope used in the disclosure of the invention
XX	
SQ	Sequence 15 AA;
Query Match	7.6%; Score 9; DB 4; Length 15;
Best Local Similarity	100.0%; Pred. No. 0.43;
Matches	9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ACMSADLEV 9
Db	2 ACMSADLEV 10
RESULT 489	
AAJ03253	
ID	AAJ03253 standard; peptide; 15 AA.
XX	
AC	AAJ03253;
XX	
DT	02-JUL-2001 (first entry)
XX	
DE	Hepatitis C virus epitope #3244.

XX PS Disclosure; Page 180; 214pp; English.

CC The present invention describes a composition comprising a prepared

CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.

CC These are derived from HCV HLA-binding motifs. They are useful in

CC vaccines for the prevention and treatment of HCV infection in humans. The

CC present sequence is an epitope used in the disclosure of the invention

XX SQ Sequence 15 AA;

Query Match 7.6%; Score 9; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.43; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 18 GGVLAALAA 26
Db 7 GGVLAALAA 15
|||||

RESULT 491

ABB77268

ID ABB77268 standard; protein; 18 AA.

AC ABB77268;

XX 28-JUN-2002 (first entry)

DT HCV bait polypeptide 16.

DE SID; selected interacting domain; HCV; hepatitis C virus; liver disease;

XX liver cancer; virucide; hepatotropic; antiinflammatory; antibacterial.

XX Hepatitis C virus strain H77.

XX EPI178116-A1.

XX 06-FEB-2002.

XX 03-AUG-2000; 2000EP-00402225.

XX 03-AUG-2000; 2000EP-00402225.

XX (HYBR-) HYBRIGENICS SA.

XX Legrain P, Whiteside S, Wojcik J;

XX WPI; 2002-208115/27.

XX N-PSDB; ABL55600.

XX New selected interacting domain polypeptides and polynucleotides, useful

PT for treating or preventing infections or pathologies caused by hepatitis

PT C virus (HCV) or those linked to HCV infection.

XX Claim 26; SEQ ID NO 92; 61pp + Sequence Listing; English.

XX The invention relates to nucleic acids encoding polypeptides which are

CC termed SID polypeptides (selected interacting domain). These polypeptides

CC are the final products of a double selection method involving a first

CC step of selection of Hepatitis C virus (HCV)-derived polynucleotides

CC through a two-hybrid system, and a second selection step involving an

CC alignment between the different polynucleotides selected at the first

CC step. The activity of polypeptides of the invention may be described as,

CC virucide, hepatotropic, antiinflammatory and antibacterial. The

CC polypeptide, polynucleotide and compositions comprising them are useful

CC for treating or preventing viral or a bacterial infection, specifically

CC infections or pathologies caused by HCV or those pathologies linked to

CC HCV infection. These may include liver disease and liver cancer. The

CC current sequence represents a HCV bait polypeptide. Note: The sequence

CC data for this patent is not represented in the specification, but is

CC based on sequence information supplied by the European Patent Office

XX SQ Sequence 18 AA;

Query Match 7.6%; Score 9; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.53; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 18 GGVLAALAA 26
Db 11 GGVLAALAA 19
|||||

RESULT 493

AAW13248

ID AAW13248 standard; peptide; 31 AA.

XX AAW13248;

XX 14-NOV-1997 (first entry)

DT

XX

Query Match 7.6%; Score 9; DB 5; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.5; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 22 AALAAAYCLS 30
Db 1 AALAAAYCLS 9
|||||

RESULT 492

AAE21994

ID AAE21994 standard; peptide; 19 AA.

XX AAE21994;

XX 25-JUL-2002 (first entry)

DT Hepatitis C virus epitope derived peptide #2.

DE Hepatitis C virus; HCV; epitope; CD4+ T-lymphocyte; infection; vaccine;

XX virucide; antiinflammatory; hepatotropic.

XX Hepatitis C virus.

XX WO200226785-A2.

XX 04-APR-2002.

XX 28-SEP-2001; 2001WO-EP011263.

XX 28-SEP-2000; 2000EP-00121138.

XX (IMMU-) IMMUSYSTEMS GMBH.

XX Gerlach JT, Diepolder H;

XX WPI; 2002-362486/39.

XX New hepatitis C virus epitopes which are specific for CD40-positive T

PT lymphocytes, useful for diagnosis and vaccination of hepatitis C virus

PT infection.

XX Claim 3; Page 14; 34pp; German.

XX The present invention relates to Hepatitis C virus (HCV) epitopes that

CC are specific for CD4+ T-lymphocytes and their derivatives with comparable

CC specificity. HCV epitopes of the invention are useful for diagnosis of

CC HCV infection (by determining the level of CD4+ cells specific for a

CC particular epitope) and particularly when formulated as a vaccine, for

CC treatment and prevention of infection. DNAs that encode HCV epitopes can

CC also be used in vaccines. The present sequence is a HCV epitope derived

CC peptide

XX SQ Sequence 19 AA;

Query Match 7.6%; Score 9; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.53; Mismatches 0; Indels 0; Gaps 0;
Matches 9; Conservative 0;

Qy 18 GGVLAALAA 26
Db 11 GGVLAALAA 19
|||||

RESULT 493

AAW13248

ID AAW13248 standard; peptide; 31 AA.

XX AAW13248;

XX 14-NOV-1997 (first entry)

DT

XX

DE Hepatitis C virus serotype 3 non-structural protein 4 antigen.
 XX HCV; serotype 3; non-structural protein 4; NS4; antigen; epitope;
 KW immunoassay; antibody; enzyme linked immunosorbent assay; ELISA;
 KW detection; discrimination.
 XX Hepatitis C virus.
 OS Hepatitis C virus.
 XX WO9708198-A1.
 XX 06-MAR-1997.
 XX 29-AUG-1996; 96WO-JP002416.
 XX 31-AUG-1995; 95JP-00223628.
 XX (SRLS-) SRL INC.
 XX Kumazawa T, Kiya Y, Tagami H;
 PI WPI; 1997-179178/16.
 XX Peptide(s) antigenic for hepatitis C virus antibodies - useful for HCV
 PT diagnosis, and serotype assigning assays.
 XX Claim 1; Page 36; 49pp; Japanese.
 CC The present peptide is a hepatitis C virus (HCV) serotype 3 non-
 CC structural protein 4 (NS4) antigen. It can be used in immunoassays for
 CC HCV antibodies, preferably an enzyme linked immunosorbent assay (ELISA).
 CC Using the peptide together with other peptides antigenic for HCV
 CC antibodies, e.g. corresponding to part of the NS4, core or NS5 region,
 CC enables different HCV serotypes to be discriminated
 XX Sequence 31 AA;
 SQ

Query Match 7.6%; Score 9; DB 2; Length 31;
 Best Local Similarity 100.0%; Pred. No. 0.79;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSCQA 68
 DB 4 DEMECSCQA 12
 |||||
 |||||

RESULT 494
 AAY67802
 ID AAY67802 standard; peptide; 35 AA.
 XX AAY67802;
 AC AAY67802;
 DT 23-MAR-2000 (first entry)
 DE Peptide #202 for detecting hepatitis C virus infection.
 XX Hepatitis C virus; HCV; increased structural stability; NS4 region;
 KW diagnostic antigen.
 KW Synthetic.
 OS WO9962945-A2.
 XX 09-DEC-1999.
 XX 04-JUN-1999; 99WO-US012446.
 XX 05-JUN-1998; 98US-0088229P.
 PR 01-SEP-1998; 98US-0098705P.
 PR 15-SEP-1998; 98US-0100422P.
 PR 28-JAN-1999; 99WO-US001726.
 XX (PEPT-) PEPTIDE SOLUTIONS INC.
 PA

PI Chowdhury MA, Bernatein D, Motesbocker MA;
 XX WPI; 2000-086953/07.
 XX Improving properties of peptides for use as diagnostic antigens or for
 PT preventing or treating infections.
 XX Claim 55; Page 73; 83pp; English.
 XX This is a peptide related to the immunoreactive region of the NS4 region
 CC of hepatitis C virus (HCV). The peptide is useful for detecting HCV
 CC infection. The invention relates to peptides derived from HCV and also
 CC HIV-1 which have been modified for use as diagnostic antigens in the
 CC treatment or prevention of infection. The structural stability of the
 CC peptides can be increased in four different ways; through the replacement
 CC of a hydrophobic amino acid with a less hydrophobic amino acid; through
 CC an increase in the amount of secondary structure (i.e. alpha helix) in
 CC the peptide; through the removal of a positive charge from the peptide,
 CC or through the constraint of the epitopic sequence via the formation of a
 CC covalent crosslink. Modified peptides of the invention are used to detect
 CC infectious agents specifically HCV. Other detectable agents include HIV-1
 CC Group O viruses; human T-cell lymphotropic virus-I or -II; and the
 CC causative agent of syphilis. The peptides can be used for prevention or
 CC treatment of infections (e.g. as vaccines, or where expressed from a
 CC transgene). More generally almost any peptide can be similarly modified,
 CC e.g. cytokines or interferons; major histocompatibility complex antigens;
 CC hormones; growth factors; tumour markers or suppressors, or antigens from
 CC many other pathogens
 XX Sequence 35 AA;
 SQ

Query Match 7.6%; Score 9; DB 3; Length 35;
 Best Local Similarity 100.0%; Pred. No. 0.88;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 IVPDKVELY 56
 DB 23 IVPDKVELY 31
 |||||
 |||||

RESULT 495
 AAW13247
 ID AAW13247 standard; peptide; 48 AA.
 XX AAW13247;
 AC AAW13247;
 DT 14-NOV-1997 (first entry)
 DE Hepatitis C virus serotype 3 non-structural protein 4 antigen.
 XX HCV; serotype 3; non-structural protein 4; NS4; antigen; epitope;
 KW immunoassay; antibody; enzyme linked immunosorbent assay; ELISA;
 KW detection; discrimination.
 XX Hepatitis C virus.
 OS WO9708198-A1.
 XX 06-MAR-1997.
 XX 29-AUG-1996; 96WO-JP002416.
 XX 31-AUG-1995; 95JP-00223628.
 XX (SRLS-) SRL INC.
 XX Kumazawa T, Kiya Y, Tagami H;
 PI WPI; 1997-179178/16.
 XX Peptide(s) antigenic for hepatitis C virus antibodies - useful for HCV
 PT diagnosis, and serotype assigning assays.
 XX

PS Claim 1; Page 36; 49pp; Japanese.

XX The present peptide is a hepatitis C virus (HCV) serotype 3 non-
 CC structural protein 4 (NS4) antigen. It can be used in immunoassays for
 CC HCV antibodies, preferably an enzyme linked immunosorbant assay (ELISA).
 CC Using the peptide together with other peptides antigenic for HCV
 CC antibodies, e.g. corresponding to part of the NS4, core or NS5 region,
 CC enables different HCV serotypes to be discriminated

XX SQ Sequence 48 AA;
 Query Match 7.6%; Score 9; DB 2; Length 48;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQA 68
 Db 21 DEMEECSQA 29
 |||||

RESULT 496
 AAR63301
 ID AAR63301 standard; protein; 119 AA.
 AC AAR63301;
 XX
 DT 25-MAR-2003 (revised)
 DT 07-AUG-1995 (first entry)
 XX
 XX Polypeptide encoded by hepatitis C virus NS3 sequence.
 XX
 XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping.
 XX
 XX Hepatitis C virus type 5a.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 14 /note= "Unspecified amino acid."
 FT Misc-difference 24 /note= "Unspecified amino acid."
 FT Misc-difference 58 /note= "Unspecified amino acid"
 FT Misc-difference 113 /note= "Unspecified amino acid"
 FT
 XX WO9425601-A2.
 XX 10-NOV-1994.
 XX
 XX 27-APR-1994; 94WO-EP001323.
 XX
 XX 27-APR-1993; 93EP-00401099.
 XX 05-AUG-1993; 93EP-00402019.
 XX
 XX (INNO-) INNOGENETICS NV SA.
 XX
 XX Maertens G, Stuyver L;
 XX WPI; 1994-358277/44.
 XX N-PSDB; AAQ78053.
 XX
 XX New polynucleotide sequences from hepatitis C virus - and related
 FT vectors, polypeptide(s) and antibodies, useful for immunisation,
 FT treatment, diagnosis and typing of HCV isolates.
 XX
 XX Disclosure; Page 150-151; 404pp; English.
 XX
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
 CC genomic sequence, a type 4 genomic sequence; or the coding region of
 CC subtype 5a, may be used as primers to amplify nucleic acid from an
 CC isolate belonging to a specific genotype, or as a probe for specific
 CC detection/classification of nucleic acid. Polypeptides encoded by the
 CC nucleotides in such compositions may be used for immunisation against
 CC HCV, for the detection of antibodies directed against HCV and for
 CC serotyping. This sequence corresponds to the NS3 region of HCV subtype 5a
 CC and is taken from a clone designated PC-1-48. (Updated on 25-MAR-2003 to
 CC correct FN field.)

CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
 CC genomic sequence, a type 4 genomic sequence; or the coding region of
 CC subtype 5a, may be used as primers to amplify nucleic acid from an
 CC isolate belonging to a specific genotype, or as a probe for specific
 CC detection/classification of nucleic acid. Polypeptides encoded by the
 CC nucleotides in such compositions may be used for immunisation against
 CC HCV, for the detection of antibodies directed against HCV and for
 CC serotyping. This sequence corresponds to the NS3 region of HCV subtype 5a
 CC and is taken from a clone designated PC-1-48. (Updated on 25-MAR-2003 to
 CC correct FN field.)

XX SQ Sequence 119 AA;
 Query Match 7.6%; Score 9; DB 2; Length 119;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
 Db 2 ACMSADLEV 10
 |||||

RESULT 497
 AAR63436
 ID AAR63436 standard; protein; 481 AA.
 XX
 AC AAR63436;
 XX
 DT 25-MAR-2003 (revised)
 DT 18-AUG-1995 (first entry)
 XX
 XX HCV polypeptide sequence.
 DE
 XX
 XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping.
 XX
 XX Hepatitis C virus.
 XX
 XX WO9425601-A2.
 XX 10-NOV-1994.
 XX
 XX 27-APR-1994; 94WO-EP001323.
 XX
 XX 27-APR-1993; 93EP-00401099.
 XX 05-AUG-1993; 93EP-00402019.
 XX
 XX (INNO-) INNOGENETICS NV SA.
 XX
 XX Maertens G, Stuyver L;
 XX WPI; 1994-358277/44.
 XX N-PSDB; AAQ78126.
 XX
 XX New polynucleotide sequences from hepatitis C virus - and related
 FT vectors, polypeptide(s) and antibodies, useful for immunisation,
 FT treatment, diagnosis and typing of HCV isolates.
 XX
 XX Disclosure; Page 291-292; 404pp; English.
 XX
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or
 CC as a probe for specific detection/classification of nucleic acid.
 CC Polypeptides encoded by the nucleotides in such compositions may be used

CC for immunisation against HCV, for the detection of antibodies directed
 CC against HCV and for serotyping. (Updated on 25-MAR-2003 to correct PN
 CC field.)

SQ Sequence 481 AA;

Query Match 7.6%; Score 9; DB 2; Length 481;
 Best Local Similarity 100.0%; Pred. No. 7.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 |||||
 Db 364 ACMSADLEV 372

RESULT 498

AAW37807
 ID AAR63377 standard; protein; 489 AA.

XX AAR63377;

XX 25-MAR-2003 (revised)

DT 18-AUG-1995 (first entry)

DE Hepatitis C virus NS3/5B polypeptide.

XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping.

OS Hepatitis C virus.

FH Key Location/Qualifiers

FT Protein 1..489

FT /note= "X's in the sequence correspond to stop codons in
 FT the corresponding nucleotide file."

XX WO9425601-A2.

XX 10-NOV-1994.

XX 27-APR-1994; 94WO-EP001323.

XX 27-APR-1993; 93EP-00401099.

PR 05-AUG-1993; 93EP-00402019.

XX (INNO-) INNOGENETICS NV SA.

XX Maertens G, Stuyver L;

XX WPI; 1994-358277/44.

DR N-PSDB; AAQ78115.

XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates.

XX Disclosure; Page 250-252; 404pp; English.

XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, may be used as primers to
 CC amplify nucleic acid from an isolate belonging to a specific genotype, or
 CC as a probe for specific detection/classification of nucleic acid.
 CC Polypeptides encoded by the nucleotides in such compositions may be used
 CC for immunisation against HCV, for the detection of antibodies directed
 CC against HCV and for serotyping. This sequence corresponds to the NS3/5B
 CC region of HCV. (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 489 AA;

Query Match 7.6%; Score 9; DB 2; Length 489;
 Best Local Similarity 100.0%; Pred. No. 8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 |||||

Db 362 ACMSADLEV 370

RESULT 499

AAW37807

ID AAW37807 standard; peptide; 615 AA.

XX AAW37807;

DT 19-AUG-1998 (first entry)

DE Nonstructural domain protein 3 of Hepatitis C virus.

XX Nonstructural protein 3; NS3; HCV; detergent-free NS3 protease;
 KW screening assay; inhibition; protease activity; diagnosis; HCV infection.

OS Hepatitis C virus.

PN WO9813482-A1.

XX 02-APR-1998.

XX 23-SEP-1997; 97WO-US017029.

PR 27-SEP-1996; 96US-0027274P.

PR 12-DEC-1996; 96GB-00025802.

XX (MERI) MERCK & CO INC.

XX Sardana VV, Blue JT;

XX WPI; 1998-230696/20.

XX Detergent-free hepatitis C virus protease NS3 - useful for screening for
 PT specific inhibitors and for diagnosing infection, is more active and
 PT stable than known enzyme preparations.

XX Disclosure; Page 14-16; 24pp; English.

XX The present sequence represents the nonstructural protein 3 (NS3) of
 CC Hepatitis C virus (HCV). The present protein is a stable, detergent-free
 CC NS3 protease. The specification describes a screening assay for compounds
 CC that inhibit NS3. The assay comprises incubating NS3 with the test
 CC compound and detecting any inhibition of protease activity that occurs.
 CC NS3 can be used to screen for inhibitory compounds and for diagnosis of
 CC HCV infection

XX Sequence 615 AA;

Query Match 7.6%; Score 9; DB 2; Length 615;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 |||||

Db 605 ACMSADLEV 613

RESULT 500

AAR82854

ID AAR82854 standard; protein; 631 AA.

XX AAR82854;

XX 25-MAR-2003 (revised)

```

DT 04-APR-1996 (first entry)
XX
DE NS3 serine protease domain.
XX
KW NS3; serine protease; hepatitis C virus; HCV; NS4A; therapy.
XX
OS Hepatitis C virus.
XX
PH Key Location/Qualifiers
FT Misc-difference 78
FT /note= "represented by Cal in the specification"
FT Misc-difference 132
FT /note= "represented by Cal in the specification"
FT Misc-difference 338
FT /note= "represented in the specification by Lgu"
FT Misc-difference 454
FT /note= "represented in the specification by Aps"
XX
PN W09522985-A1.
XX
XX 31-AUG-1995.
XX
XX 14-FEB-1995; 95WO-IT000018.
XX
XX 23-FEB-1994; 94IT-RM000092.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Failla C, Tomei L;
XX
XX WPI; 1995-311381/40.
XX
XX In vitro reproduction of hepatitis C virus NS3 protease activity - by
FT including the NS4A cofactor in the mixt., useful for screening cpds. that
FT inhibit NS3.
XX
XX Claim 6; Page 16-18; 26pp; English.
XX
XX This sequence represents the Hepatitis C virus (HCV) NS3 serine protease
CC domain. The NS3 serine protease domain requires NS4A (see AAR82855) as a
CC cofactor. Optimal serine protease activity is obtained when NS4A and NS3
CC are present in a 1:1 ratio. The cleavage site between these two proteins
CC on the HCV genome can be mutated so that the components remain covalently
CC bonded. These sequences are included in a composition that can be used in
CC an assay system. This assay system can be used to select compounds that
CC inhibit NS3 activity, e.g. potential therapeutic agents. NS4A can be used
CC for screening enzyme inhibitors. (Updated on 25-MAR-2003 to correct PR
CC field.)
XX
SQ Sequence 631 AA;
Query Match 7.6%; Score 9; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACMSADLEV 9
Db 621 ACMSADLEV 629
RESULT 501
AAW26160
ID AAW26160 standard; peptide; 631 AA.
XX
XX AAW26160;
AC
XX
XX 24-NOV-1997 (first entry)
DT
DE Serine protease NS3 region.
XX
XX Serine protease; NS3 region; HCV; hepatitis C virus; monoclonal antibody;
KW Igg; immunocyte; myeloma cell; inhibitor; neutralising agent.
XX
XX
OS Hepatitis C virus.
XX
PN JP09206076-A.
XX
XX 12-AUG-1997.
XX
XX 06-FEB-1996; 96JP-00020321.
XX
XX 06-FEB-1996; 96JP-00020321.
XX
XX (NIHA ) JAPAN ENERGY CORP.
XX
XX WPI; 1997-451976/42.
XX
XX N-PSDB; AAT80095.
XX
XX Monoclonal antibody against hepatitis C virus serine protease - useful
FT for diagnosis of HCV infection.
XX
XX Claim 2; Page 14-17; 22pp; Japanese.
XX
XX This sequence represents the NS3 region of the hepatitis C virus (HCV)
CC serine protease. The fragment of this sequence represented in AAW26159 is
CC recognised by the monoclonal antibody of the invention. The monoclonal
CC antibody belongs to the IgG class of antibodies. The antibody is produced
CC by fusion cells formed between immunocytes from a HCV infected mammal,
CC and myeloma cells from the same mammal. The antibody is used for
CC diagnosis of HCV infection. As this monoclonal antibody can bind to a
CC certain chain in serine protease and inhibits the enzyme specifically, it
CC can also be used as an agent for neutralising the activity of serine
CC protease
XX
SQ Sequence 631 AA;
Query Match 7.6%; Score 9; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACMSADLEV 9
Db 621 ACMSADLEV 629
RESULT 502
AAW14354
ID AAY14354 standard; protein; 631 AA.
XX
XX AAY14354;
AC
XX
XX 17-AUG-1999 (first entry)
DT
XX
XX Hepatitis C virus Ser/Thr protease amino acid sequence.
DE
XX
XX Complementarity determining region; CDR; monoclonal antibody; Mab;
KW hepatitis C virus; HCV; protease.
XX
XX Hepatitis C virus.
OS
XX
XX JP11127861-A.
XX
XX 18-MAY-1999.
XX
XX 29-OCT-1997; 97JP-00297451.
XX
XX 29-OCT-1997; 97JP-00297451.
XX
XX (NIHA ) JAPAN ENERGY CORP.
XX
XX WPI; 1999-350322/30.
XX
XX N-PSDB; AAX57785.
XX
XX Neutralized antibody partial peptide derived from hepatitis C virus -
PT useful for inhibiting Hepatitis C Virus (HCV) serine protease activity.
XX

```

PS Disclosure; Page 14-16; 32pp; Japanese.

XX This sequence represents the amino acid sequence of a Ser/Thr protease
 CC from hepatitis C virus. The invention relates to the use of partial
 CC peptides (AAV14348-Y14353) from the anti-HCV neutralising Mab 8D4 for
 CC inhibiting HCV serine protease activity
 XX
 XX Sequence 631 AA;

Query Match 7.6%; Score 9; DB 2; Length 631;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
 |||||
 Db 621 ACMSADLEV 629

RESULT 503
 AAY15806
 ID AAY15806 standard; protein; 631 AA.
 XX
 AC AAY15806;
 XX
 XX 27-JUL-1999 (first entry)
 DT
 XX HCV strain J antigen sequence.
 DE
 XX
 XX Antigen peptide; HCV J strain; chain peptide; detection; antibody.
 KW
 XX
 XX Hepatitis C virus.
 OS
 XX JP11124398-A.
 FN
 XX
 XX 11-MAY-1999.
 PD
 XX
 XX 22-OCT-1997; 97JP-00290165.
 PF
 XX
 XX 22-OCT-1997; 97JP-00290165.
 PR
 XX
 XX (NIIHA) JAPAN ENERGY CORP.
 PA
 XX
 XX WPI; 1999-341639/29.
 DR
 XX N-PSDB; AAX59785.
 DR
 XX
 XX New antigen peptide from hepatitis C virus - useful in examination agent
 PT for antibody.
 PT
 XX
 XX Disclosure; Page 10-13; 25pp; Japanese.

XX The specification describes an antigen peptide derived from hepatitis C
 CC virus (HCV), which comprises a chain peptide of at least 12 amino acid
 CC residues. The peptide chain comprises at least one amino acid sequence
 CC selected from Gly Trp Pro or AAY15764-66. The peptide is useful in
 CC methods to detect antibodies directed against HCV. The present sequence
 CC represents a HCV strain J antigen
 XX
 XX Sequence 631 AA;

Query Match 7.6%; Score 9; DB 2; Length 631;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
 |||||
 Db 621 ACMSADLEV 629

RESULT 504
 ADL18158
 ID ADL18158 standard; protein; 631 AA.
 XX
 AC ADL18158;

XX 06-MAY-2004 (first entry)
 DT
 XX Hepatitis C virus NS3 protease protein SEQ ID NO:78.
 DE
 XX chimeric protein; signal protein; trafficking signal targeting;
 KW proteolytic cleavage site; protease; protease inhibitor; enzyme.
 XX
 OS Hepatitis C virus.
 XX
 XX WO2003014381-A1.
 PN
 XX 20-FEB-2003.
 PD
 XX
 XX 08-AUG-2002; 2002WO-KR001515.
 PF
 XX
 XX 10-AUG-2001; 2001KR-00048123.
 PR
 XX (AHRA-) AHRA BIOSYSTEMS INC.
 PA
 XX
 XX Hwang I, Kim DH, Lee YJ;
 PI
 XX
 XX WPI; 2003-256596/25.
 DR
 XX N-PSDB; ADL18157.
 DR
 XX
 XX New chimeric protein, useful for detecting protease inhibitors inside the
 PT cell or tissue.
 PT
 XX
 XX Disclosure; SEQ ID NO 78; 214pp; English.

XX The present invention describes a chimeric protein comprising at least
 CC one signal protein that has a trafficking signal targeting to a
 CC subcellular organelle and at least one proteolytic cleavage site for a
 CC protease. The chimeric protein is constructed, so that: (a) the
 CC trafficking signals of all the signal proteins are inactivated by linking
 CC the proteolytic site or a signal masking protein through the proteolytic
 CC site to the N- or C- terminus of the signal proteins, and so the chimeric
 CC protein is present in cytosol; (b) the trafficking signal of at least one
 CC signal protein is activated when the proteolytic cleavage site is cleaved
 CC by the protease, and as a result at least one fragment protein that
 CC includes the activated signal protein is transported to a subcellular
 CC organelle; and (c) the chimeric protein is labelled with at least one
 CC fluorescent protein and the position and intensity distribution of the
 CC fluorescent label signal in the cell is altered depending on the cleavage
 CC by the protease. Also described: (1) a recombinant gene comprising a
 CC nucleic acid sequence encoding the chimeric protein which is constructed
 CC to express the chimeric protein in a cell; (2) a cell transformed with
 CC the recombinant gene or vector; (3) analysing the activity of a protease
 CC in vivo; (4) screening protease inhibitors in vivo; (5) a system for
 CC detecting a protease inside a cell; (6) a nucleic acid comprising the
 CC sequence encoding the chimeric protein for detecting protease activity in
 CC a cell; (7) a vector comprising the nucleic acid; (8) a kit for detecting
 CC a protease inside a cell comprising the chimeric protein or the vector;
 CC (9) detecting a protease inside a cell or tissue; and (10) detecting a
 CC protease inhibitor in vivo. The chimeric protein is useful for detecting
 CC protease inhibitors inside the cell or tissue. The present sequence
 CC represents HCV NS3 protease, which is used in the exemplification of the
 CC present invention.
 XX
 XX Sequence 631 AA;

Query Match 7.6%; Score 9; DB 7; Length 631;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
 |||||
 Db 621 ACMSADLEV 629

RESULT 505
 AAY17897
 ID AAY17897 standard; protein; 632 AA.

PI Brett S, Hamblin PA, Ogilvie L;
 DR WPI; 2004-420614/39.
 DR N-PSDB; ADO79398.
 XX
 PT New Hepatitis C virus (HCV) vaccine having a polynucleotide that encodes
 PT the polypeptide sequences of the HCV core, NS3, NS4B and NS5B proteins,
 PT for use in medicine, in particular for manufacturing a medicament for the
 PT treatment of HCV.
 XX
 XX
 PS Example 2; Page 26-27; 79pp; English.
 CC The present sequence is that of the NS3 protein of hepatitis C virus
 CC (HCV). The sequence is the translation sequence of a polynucleotide in
 CC which codon usage was altered to resemble that of highly expressed human
 CC genes. HCV vaccines of the invention comprise a polynucleotide that
 CC encodes the HCV proteins Core, NS3, NS4B and NS5B, and does not encode
 CC the NS4A and/or NS5A proteins. The proteins may be expressed as
 CC individual proteins or as fusion proteins. Preferred fusions include
 CC double fusions between NS4B and NS5B and between Core and NS3. The
 CC vaccines are useful for the treatment or prevention of an HCV infection.
 XX
 XX
 SQ Sequence 632 AA;
 Query Match 7.6%; Score 9; DB 8; Length 632;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACMSADLEV 9
 Db 622 ACMSADLEV 630
 |||||
 RESULT 508
 AAY17894
 ID AAY17894 standard; protein; 646 AA.
 XX
 AC AAY17894;
 XX
 DT 07-SEP-1999 (first entry)
 XX
 DE HCV NS4A-NS3 complex SEQ ID NO:72.
 XX
 KW HCV; hepatitis C virus; single chain recombinant complex; linker;
 KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
 KW hydrophobic domain; covalent complex; detection; inhibitor.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FN WO9928482-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 24-NOV-1998; 98WO-US024528.
 XX
 PR 28-NOV-1997; 97US-0067315P.
 PR 28-JUL-1998; 98US-0094331P.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Malcolm BA, Taremi SS, Weber PC, Yao N;
 XX
 DR WPI; 1999-385385/32.
 XX
 PD 10-JUN-1999.
 XX
 PF 24-NOV-1998; 98WO-US024528.
 XX
 PR 28-NOV-1997; 97US-0067315P.
 PR 28-JUL-1998; 98US-0094331P.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Malcolm BA, Taremi SS, Weber PC, Yao N;
 XX
 DR WPI; 1999-385385/32.
 XX
 PT New hepatitis C virus covalent complexes.
 XX
 PS Example 2; Page 135-137; 211pp; English.
 CC The present invention describes a covalent hepatitis C virus (HCV) NS4A-
 CC NS3 complex comprising a central hydrophobic domain of native HCV NS4A
 CC peptide, a linker, and an HCV NS3 serine protease domain, where the
 CC hydrophobic domain of native HCV NS4A peptide is tethered by the linker
 CC to the amino terminus of the HCV NS3 protease domain.

CC to the amino terminus of the HCV NS3 protease domain. The present
 CC sequence represents an example of the above complex. The covalent NS4A-
 CC NS3 complexes are useful for structural determination and determination
 CC of mode of binding of HCV inhibitors by NMR spectroscopy. They can also
 CC be used for detecting inhibitors of the protease activity. The helicase
 CC activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes
 CC are more soluble, stable and active than the non- covalent protease-
 CC peptide complexes previously available
 XX
 SQ Sequence 646 AA;
 Query Match 7.6%; Score 9; DB 2; Length 646;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACMSADLEV 9
 Db 636 ACMSADLEV 644
 |||||
 RESULT 509
 AAY17892
 ID AAY17892 standard; protein; 646 AA.
 XX
 AC AAY17892;
 XX
 DT 07-SEP-1999 (first entry)
 XX
 DE HCV NS4A-NS3 complex SEQ ID NO:66.
 XX
 KW HCV; hepatitis C virus; single chain recombinant complex; linker;
 KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
 KW hydrophobic domain; covalent complex; detection; inhibitor.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FN WO9928482-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 24-NOV-1998; 98WO-US024528.
 XX
 PR 28-NOV-1997; 97US-0067315P.
 PR 28-JUL-1998; 98US-0094331P.
 XX
 PA (SCHE) SCHERING CORP.
 XX
 PI Malcolm BA, Taremi SS, Weber PC, Yao N;
 XX
 DR WPI; 1999-385385/32.
 XX
 PD New hepatitis C virus covalent complexes.
 XX
 PS Example 2; Page 129-131; 211pp; English.
 CC The present invention describes a covalent hepatitis C virus (HCV) NS4A-
 CC NS3 complex comprising a central hydrophobic domain of native HCV NS4A
 CC peptide, a linker, and an HCV NS3 serine protease domain, where the
 CC hydrophobic domain of native HCV NS4A peptide is tethered by the linker
 CC to the amino terminus of the HCV NS3 protease domain. The present
 CC sequence represents an example of the above complex. The covalent NS4A-
 CC NS3 complexes are useful for structural determination and determination
 CC of mode of binding of HCV inhibitors by NMR spectroscopy. They can also
 CC be used for detecting inhibitors of the protease activity, the helicase
 CC activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes
 CC are more soluble, stable and active than the non- covalent protease-
 CC peptide complexes previously available
 XX
 SQ Sequence 646 AA;
 Query Match 7.6%; Score 9; DB 2; Length 646;
 Best Local Similarity 100.0%; Pred. No. 10;

```
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 510
AAY24950
ID AAY24950 standard; protein; 646 AA.
XX
AC AAY24950;
XX
DT 07-SEP-1999 (first entry)
DE HCV NS4A-NS3 complex SEQ ID NO:63.
XX
KW HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
PN WO9928482-A2.
XX
DE 10-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US024528.
XX
PR 28-NOV-1997; 97US-0067315P.
PR 28-JUL-1998; 98US-0094331P.
XX
PA (SCHE ) SCHERING CORP.
XX
PI Malcolm BA, Taremi SS, Weber PC, Yao N;
XX
DR WPI; 1999-385385/32.
XX
PT New hepatitis C virus covalent complexes.
XX
PS Example 2; Page 126-128; 21lpp; English.
XX
CC The present invention describes a covalent hepatitis C virus (HCV) NS4A-
CC NS3 complex comprising a central hydrophobic domain of native HCV NS4A
CC peptide, a linker, and an HCV NS3 serine protease domain, where the
CC hydrophobic domain of native HCV NS4A peptide is tethered by the linker
CC to the amino terminus of the HCV NS3 protease domain. The present
CC sequence represents an example of the above complex. The covalent NS4A-
CC NS3 complexes are useful for structural determination and determination
CC of mode of binding of HCV inhibitors by NMR spectroscopy. They can also
CC be used for detecting inhibitors of the protease activity, the helicase
CC activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes
CC are more soluble, stable and active than the non- covalent protease-
CC peptide complexes previously available
XX
SQ Sequence 646 AA;

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 511
AAY24942
ID AAY24942 standard; protein; 665 AA.
XX
AC AAY24942;
XX
```

```
DT 07-SEP-1999 (first entry)
XX HCV NS4A-NS3 complex SEQ ID NO:13.
XX
KW HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
PN WO9928482-A2.
XX
PD 10-JUN-1999.
XX
PF 24-NOV-1998; 98WO-US024528.
XX
PR 28-NOV-1997; 97US-0067315P.
PR 28-JUL-1998; 98US-0094331P.
XX
PA (SCHE ) SCHERING CORP.
XX
PI Malcolm BA, Taremi SS, Weber PC, Yao N;
XX
DR WPI; 1999-385385/32.
XX
PT New hepatitis C virus covalent complexes.
XX
PS Claim 6; Page 88-90; 21lpp; English.
XX
CC The present invention describes a covalent hepatitis C virus (HCV) NS4A-
CC NS3 complex comprising a central hydrophobic domain of native HCV NS4A
CC peptide, a linker, and an HCV NS3 serine protease domain, where the
CC hydrophobic domain of native HCV NS4A peptide is tethered by the linker
CC to the amino terminus of the HCV NS3 protease domain. The present
CC sequence represents a specifically claimed example of the above complex.
CC The covalent NS4A-NS3 complexes are useful for structural determination
CC and determination of mode of binding of HCV inhibitors by NMR
CC spectroscopy. They can also be used for detecting inhibitors of the
CC protease activity, the helicase activity and the ATPase activity of NS3.
CC The covalent NS4A-NS3 complexes are more soluble, stable and active than
CC the non-covalent protease-peptide complexes previously available
XX
SQ Sequence 665 AA;

Query Match 7.6%; Score 9; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664

RESULT 512
AAY24941
ID AAY24941 standard; protein; 665 AA.
XX
AC AAY24941;
XX
DT 07-SEP-1999 (first entry)
XX
DE HCV NS4A-NS3 complex SEQ ID NO:12.
XX
KW HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
PN WO9928482-A2.
XX
```


CC protease activity, the helicase activity and the ATPase activity of NS3.
 CC The covalent NS4A-NS3 complexes are more soluble, stable and active than
 CC the non-covalent protease-peptide complexes previously available
 XX
 SQ Sequence 665 AA;

Query Match 7.6%; Score 9; DB 2; Length 665;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 |||||
 Db 656 ACMSADLEV 664

RESULT 515
 AAY24940
 ID AAY24940 standard; protein; 665 AA.
 XX
 AC AAY24940;

XX
 DT 07-SEP-1999 (first entry)
 XX
 DE HCV NS4A-NS3 complex SEQ ID NO:11.
 XX

KW HCV; hepatitis C virus; single chain recombinant complex; linker;
 KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
 KW hydrophobic domain; covalent complex; detection; inhibitor.

XX
 OS Hepatitis C virus.
 OS Synthetic.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

XX 28-JUL-1998; 98US-0094331P.

XX (SCHE) SCHERING CORP.

XX Malcolm BA, Taremi SS, Weber PC, Yao N;

XX WPI; 1999-385385/32.

XX New hepatitis C virus covalent complexes.

XX Claim 6; Page 83-85; 21pp; English.

XX The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available

XX Sequence 665 AA;

Query Match 7.6%; Score 9; DB 2; Length 665;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 |||||
 Db 656 ACMSADLEV 664

RESULT 516
 AAY24943

XX AAY24943 standard; protein; 665 AA.

XX AAY24943;

XX 07-SEP-1999 (first entry)

XX HCV NS4A-NS3 complex SEQ ID NO:14.

XX HCV; hepatitis C virus; single chain recombinant complex; linker;
 KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
 KW hydrophobic domain; covalent complex; detection; inhibitor.

XX Hepatitis C virus.

XX Synthetic.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

XX 28-JUL-1998; 98US-0094331P.

XX (SCHE) SCHERING CORP.

XX Malcolm BA, Taremi SS, Weber PC, Yao N;

XX WPI; 1999-385385/32.

XX New hepatitis C virus covalent complexes.

XX Claim 6; Page 90-92; 21pp; English.

XX The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available

XX Sequence 665 AA;

Query Match 7.6%; Score 9; DB 2; Length 665;
 Best Local Similarity 100.0%; Pred. No. 10;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 |||||
 Db 656 ACMSADLEV 664

RESULT 517
 AAY24947

XX AAY24947 standard; protein; 665 AA.

XX AAY24947;

XX 07-SEP-1999 (first entry)

XX HCV NS4A-NS3 complex SEQ ID NO:18.

XX HCV; hepatitis C virus; single chain recombinant complex; linker;

KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.

OS Hepatitis C virus.
OS Synthetic.

PN WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

XX 28-JUL-1998; 98US-0094331P.

XX (SCHE) SCHERING CORP.

XX Malcolm BA, Taremi SS, Weber PC, Yao N;

XX WPI; 1999-385385/32.

XX New hepatitis C virus covalent complexes.

XX Claim 6; Page 100-102; 21lpp; English.

XX The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available

XX Sequence 665 AA;

Query Match 7.6%; Score 9; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9

Db 656 ACMSADLEV 664

RESULT 518

AAV24946

ID AAV24946 standard; protein; 665 AA.

XX AC AAV24946;

XX 07-SEP-1999 (first entry)

XX HCV NS4A-NS3 complex SEQ ID NO:17.

XX HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.

OS Hepatitis C virus.

OS Synthetic.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

PR 28-JUL-1998; 98US-0094331P.
XX (SCHE) SCHERING CORP.
XX Malcolm BA, Taremi SS, Weber PC, Yao N;
XX WPI; 1999-385385/32.
XX New hepatitis C virus covalent complexes.
XX Claim 6; Page 97-99; 21lpp; English.
XX The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available

XX Sequence 665 AA;
Query Match 7.6%; Score 9; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACMSADLEV 9
Db 656 ACMSADLEV 664
RESULT 519
ADA12164
ID ADA12164 standard; protein; 665 AA.
XX AC ADA12164;
XX 06-NOV-2003 (first entry)
XX Hepatitis C virus H1S-NS4A_21-32_-GSGS-NS3_3-631 protein.
XX protein co-ordinate data; HCV; hepatitis C virus; X-ray diffraction;
KW neutron diffraction; structure-based design.
XX Hepatitis C virus.
OS US6524589-B1.
XX 25-FEB-2003.
XX 05-APR-2000; 2000US-00543376.
XX 08-APR-1999; 99US-0128253P.
XX (SCHE) SCHERING CORP.
XX Reichert P, Prosise WW, Taremi SS, Yao N, Weber PC;
XX WPI; 2003-605335/57.
XX Crystalline composition comprising hepatitis C virus (HCV) NS3/NS4A polypeptide complex, useful for determining three-dimensional structure of HCV NS3/NS4A complex, and modeling tertiary structure of related proteins.
XX Claim 1; Col 151-154; 84pp; English.
XX The invention relates to a crystalline composition comprising a hepatitis

CC C virus (HCV) NS3/NS4A polypeptide complex. The composition has
CC therapeutic applications, and is useful to obtain a determination of the
CC three-dimensional structure of HCV NS3/NS4A complex to high resolution,
CC and for modeling the tertiary structure of related proteins and/or
CC protein complexes. The composition is useful in X-ray or neutron
CC diffraction analysis to determine the three-dimensional structure of
CC NS3/NS4A polypeptide complex and in particular to assist in the
CC identification of the protein's active and effector sites. Knowledge of
CC these sites and solvent accessible residues allow structure-based design
CC and construction of agonists and antagonists for NS3 and/or NS3/NS4A
CC polypeptide complex. The present sequence represents the amino acid
CC sequence of the hepatitis C virus HIS-NS4A_21-32_-GSGS-NS3_3-631 protein.
XX
SQ Sequence 665 AA;

Query Match 7.6%; Score 9; DB 6; Length 665;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664
|||||

RESULT 520

ADA12165
ID ADA12165 standard; protein; 665 AA.

AC ADA12165;

DT 06-NOV-2003 (first entry)

XX Hepatitis C virus HIS-NS4A_21-32_-GSGS-NS3_3-631/SL139A protein.

DE protein co-ordinate data; HCV; hepatitis C virus; X-ray diffraction;
KW neutron diffraction; structure-based design.

OS Hepatitis C virus.

PN US6524589-B1.

XX 25-FEB-2003.

XX 05-APR-2000; 2000US-00543376.

XX 08-APR-1999; 99US-0128253P.

XX (SCHE) SCHERING CORP.

PI Reichert P, Proise WW, Taremi SS, Yao N, Weber PC;

XX WPI; 2003-605335/57.

XX Crystalline composition comprising hepatitis C virus (HCV) NS3/NS4A
PT polypeptide complex, useful for determining three-dimensional structure
PT of HCV NS3/NS4A complex, and modeling tertiary structure of related
PT proteins.

XX Claim 1; Col 153-158; 84pp; English.

CC The invention relates to a crystalline composition comprising a hepatitis
CC C virus (HCV) NS3/NS4A polypeptide complex. The composition has
CC therapeutic applications, and is useful to obtain a determination of the
CC three-dimensional structure of HCV NS3/NS4A complex to high resolution,
CC and for modeling the tertiary structure of related proteins and/or
CC protein complexes. The composition is useful in X-ray or neutron
CC diffraction analysis to determine the three-dimensional structure of
CC NS3/NS4A polypeptide complex and in particular to assist in the
CC identification of the protein's active and effector sites. Knowledge of
CC these sites and solvent accessible residues allow structure-based design
CC and construction of agonists and antagonists for NS3 and/or NS3/NS4A
CC polypeptide complex. The present sequence represents the amino acid
CC sequence of the hepatitis C virus HIS-NS4A_21-32_-GSGS-NS3_3-631/SL139A

CC protein.
XX
SQ Sequence 665 AA;

Query Match 7.6%; Score 9; DB 6; Length 665;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664
|||||

RESULT 521

ADA12166
ID ADA12166 standard; protein; 665 AA.

AC ADA12166;

DT 06-NOV-2003 (first entry)

XX Hepatitis C virus HIS-NS4A_21-32_-GSGS-NS3_3-631/SL139A R180A protein.
DE protein co-ordinate data; HCV; hepatitis C virus; X-ray diffraction;
KW neutron diffraction; structure-based design.

OS Hepatitis C virus.

PN US6524589-B1.

XX 25-FEB-2003.

XX 05-APR-2000; 2000US-00543376.

XX 08-APR-1999; 99US-0128253P.

XX (SCHE) SCHERING CORP.

PI Reichert P, Proise WW, Taremi SS, Yao N, Weber PC;

XX WPI; 2003-605335/57.

XX Crystalline composition comprising hepatitis C virus (HCV) NS3/NS4A
PT polypeptide complex, useful for determining three-dimensional structure
PT of HCV NS3/NS4A complex, and modeling tertiary structure of related
PT proteins.

XX Claim 1; Col 157-162; 84pp; English.

CC The invention relates to a crystalline composition comprising a hepatitis
CC C virus (HCV) NS3/NS4A polypeptide complex. The composition has
CC therapeutic applications, and is useful to obtain a determination of the
CC three-dimensional structure of HCV NS3/NS4A complex to high resolution,
CC and for modeling the tertiary structure of related proteins and/or
CC protein complexes. The composition is useful in X-ray or neutron
CC diffraction analysis to determine the three-dimensional structure of
CC NS3/NS4A polypeptide complex and in particular to assist in the
CC identification of the protein's active and effector sites. Knowledge of
CC these sites and solvent accessible residues allow structure-based design
CC and construction of agonists and antagonists for NS3 and/or NS3/NS4A
CC polypeptide complex. The present sequence represents the amino acid
CC sequence of the hepatitis C virus HIS-NS4A_21-32_-GSGS-NS3_3-631/SL139A
CC R180A protein.

SQ Sequence 665 AA;

Query Match 7.6%; Score 9; DB 6; Length 665;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664
|||||

RESULT 522

AAV17893
ID AAY17893 standard; protein; 667 AA.
XX
AC AAY17893;
XX
DT 07-SEP-1999 (first entry)
XX
XX HCV NS4A-NS3 complex SEQ ID NO:69.
XX
XX HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX WO9928482-A2.
XX
XX 10-JUN-1999.
XX
XX 24-NOV-1998; 98WO-US024528.
XX
XX HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX WO9928482-A2.
XX
XX 10-JUN-1999.
XX
XX 24-NOV-1998; 98WO-US024528.
XX
XX 28-NOV-1997; 97US-0067315P.
PR 28-JUL-1998; 98US-0094331P.
XX
PA (SCHE) SCHERING CORP.
XX
PI Malcolm BA, Taremi SS, Weber PC, Yao N;
XX
XX WPI; 1999-385385/32.
XX
XX New hepatitis C virus covalent complexes.
XX
XX Example 2; Page 132-134; 21lpp; English.

The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity. They can also activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available

SQ Sequence 667 AA;

Query Match 7.6%; Score 9; DB 2; Length 667;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9

Db 636 ACMSADLEV 644

RESULT 523

AAV17891
ID AAY17891 standard; protein; 667 AA.
XX
AC AAY17891;
XX
DT 07-SEP-1999 (first entry)
XX
XX HCV NS4A-NS3 complex SEQ ID NO:60.
XX
XX HCV; hepatitis C virus; single chain recombinant complex; linker;

KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.

OS Hepatitis C virus.
OS Synthetic.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

PR 28-JUL-1998; 98US-0094331P.

XX (SCHE) SCHERING CORP.

XX PI Malcolm BA, Taremi SS, Weber PC, Yao N;

XX WPI; 1999-385385/32.

XX New hepatitis C virus covalent complexes.

XX Example 2; Page 123-126; 21lpp; English.

The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity. They can also activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available

SQ Sequence 667 AA;

Query Match 7.6%; Score 9; DB 2; Length 667;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9

Db 636 ACMSADLEV 644

RESULT 524

AAV24949
ID AAY24949 standard; protein; 671 AA.

XX AAY24949;

XX 07-SEP-1999 (first entry)

XX HCV NS4A-NS3 complex SEQ ID NO:20.

KW HCV; hepatitis C virus; single chain recombinant complex; linker;
KW NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
KW hydrophobic domain; covalent complex; detection; inhibitor.

XX Hepatitis C virus.

OS Synthetic.

XX WO9928482-A2.

XX 10-JUN-1999.

XX 24-NOV-1998; 98WO-US024528.

XX 28-NOV-1997; 97US-0067315P.

CC virus antigen HCV NS3.
XX Sequence 798 AA;
SQ

Query Match 7.6%; Score 9; DB 8; Length 798;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 788 ACMSADLEV 796
|||||

RESULT 527
ADX40799
ID ADX40799 standard; protein; 3014 AA.
XX
AC ADX40799;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #22.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX

Query Match 7.6%; Score 9; DB 9; Length 3014;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 1648 ACMSADLEV 1656
|||||

RESULT 528
ADX40821
ID ADX40821 standard; protein; 3014 AA.
XX
AC ADX40821;

Query Match 7.6%; Score 9; DB 9; Length 3014;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 1648 ACMSADLEV 1656
|||||

RESULT 529
AAJ02221
ID AAJ02221 standard; peptide; 8 AA.
XX
AC AAJ02221;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #2212.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX

XX 21-APR-2005 (first entry)
XX HCV polymerase protein #44.
XX Immune stimulation; polymerase; enzyme.
XX Hepatitis C virus.
XX WO2005012502-A2.
XX 10-FEB-2005.
XX 29-MAR-2004; 2004WO-US009510.
XX 28-MAR-2003; 2003US-0458026P.
XX (EPIM-) EPIMUNE INC.
XX Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.
XX Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.
XX Disclosure; Page 388-440; 458pp; English.
XX The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.
XX Sequence 3014 AA;
SQ

Query Match 7.6%; Score 9; DB 9; Length 3014;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 1648 ACMSADLEV 1656
|||||

RESULT 529
AAJ02221
ID AAJ02221 standard; peptide; 8 AA.
XX
AC AAJ02221;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #2212.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
XX 19-JUL-2000; 2000WO-US019774.
PF
XX 19-JUL-1999; 99US-00357737.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT
XX Disclosure; Page 105; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 8 AA;
Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22 AALAAAYCL 29
Db 1 AALAAAYCL 8
RESULT 536
AAJ01613
ID AAJ01613 standard; peptide; 8 AA.
AC
AC AAJ01613;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1604.
XX
DE Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
XX WO200121189-A1.
PN
XX 29-MAR-2001.
PD
XX 19-JUL-2000; 2000WO-US019774.
PF
XX 19-JUL-1999; 99US-00357737.
PR
XX (EPIM-) EPIMMUNE INC.
PA
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT

XX Disclosure; Page 141; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 8 AA;
Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 1 CMSADLEV 8
RESULT 537
AAJ01061
ID AAJ01061 standard; peptide; 8 AA.
XX
XX AAJ01061;
AC
XX 02-JUL-2001 (first entry)
DT
XX Hepatitis C virus epitope #1052.
DE
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
PR
XX 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMMUNE INC.
PA
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
PT
XX Disclosure; Page 127; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 8 AA;
Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALAAAY 27
Db 1 VLAALAAAY 8
RESULT 538
AAJ00206

```

ID AAJ00206 standard; peptide; 8 AA.
XX
AC AAJ00206;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #197.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
A new composition useful as a vaccines against hepatitis C virus.
Disclosure; Page 105; 214pp; English.
XX
The present invention describes a composition comprising a prepared
hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 8 AA;
Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 20 VLAAALAA 27
DB 1 VLAAALAA 8
RESULT 540
AAJ01305
ID AAJ01305 standard; peptide; 8 AA.
XX
AC AAJ01305;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1296.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
A new composition useful as a vaccines against hepatitis C virus.
Disclosure; Page 134; 214pp; English.
XX
The present invention describes a composition comprising a prepared
hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 8 AA;
Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 CMSADLEV 9
DB 1 CMSADLEV 8
RESULT 539
AAJ00156
ID AAJ00156 standard; peptide; 8 AA.
XX
AC AAJ00156;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #147.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 2 CMSADLEV 9
DB 1 CMSADLEV 8
RESULT 539
AAJ00156
ID AAJ00156 standard; peptide; 8 AA.
XX
AC AAJ00156;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #147.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
```

```

Qy 22 AALAAAYCL 29
Db 1 AALAAAYCL 8

RESULT 541
AAJ01715
ID AAJ01715 standard; peptide; 8 AA.
XX AC AAJ01715;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1706.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 143; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 8 AA;

Query Match 6.8%; Score 8; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAAAY 27
Db 1 VLAALAAAY 8

RESULT 542
AAR59164
ID AAR59164 standard; peptide; 9 AA.
XX AC AAR59164;
XX DT 25-MAR-2003 (revised)
XX DT 03-MAY-1995 (first entry)
XX DE Peptide fragment (1.0131) of HCV binds HLA-A2.1.
XX KW antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
XX KW core antigen; surface antigen; pharmaceutical composition; in vivo;
XX KW ex vivo; therapeutic; diagnostic; MHC class I molecule;
XX KW major histocompatibility complex; HLA-A2.1; 9mer; 10mer; anchor;
XX KW human leukocyte antigen.
XX

OS Hepatitis C virus.
XX PN WO9420127-A1.
XX PD 15-SEP-1994.
XX PF 04-MAR-1994; 94WO-US002353.
XX PR 05-MAR-1993; 93US-00027146.
XX PR 04-JUN-1993; 93US-00073205.
XX PR 29-NOV-1993; 93US-00159184.
XX PA (CYTE-) CYTEL CORP.
XX PI Grey HM, Sette A, Sidney J, Kast W;
XX PI WPI; 1994-302678/37.
XX DR
XX PT Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
XX PT treatment or prophylaxis of cancer, virus infection or autoimmune
XX PT diseases.
XX PS Example 5; Page 102; 138pp; English.
XX CC AAR59104-264 are immunogenic 9mer peptides that contain a HLA-A2.1
XX CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity
XX CC of at least 1% as compared to a reference peptide (AAR71293). AAR59164
XX CC has an IC50 of 0.0067 and the sequence occurs at position 1180 in the HCV
XX CC LORF protein. The peptides of the invention can induce cytotoxic T
XX CC lymphocytes which can react with target cells. They can be used for the
XX CC treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 9 AA;

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 1 CMSADLEV 8

RESULT 543
AAU26650
ID AAU26650 standard; peptide; 9 AA.
XX AC AAU26650;
XX DT 18-DEC-2001 (first entry)
XX DE Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #93.
XX KW Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
XX KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
XX KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
XX KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
XX OS Homo sapiens.
XX PN WO200162776-A1.
XX PD 30-AUG-2001.
XX PF 23-FEB-2000; 2000WO-US004655.
XX PR 23-FEB-2000; 2000WO-US004655.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Kast WM, Southwood S;
XX

```

DR WPI; 2001-582039/65.
XX Composition for treating viral diseases and cancer comprises an
PT immunogenic peptide having an HLA-A2.1 binding motif.
XX
XX Example 1; Page 30; 85pp; English.
XX
XX Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
CC the invention are capable of specifically binding glycoproteins encoded
CC by HLA alleles and inducing a cytotoxic T cell response against an
CC antigen in a patient expressing HLA-A2.1. This method is useful for the
CC treatment, prevention and diagnosis of pathological states such as viral
CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
CC acuminatum. The peptides are used for treatment of chronic infection and
CC for stimulating the immune system to eliminate virus-infected cells
XX
SQ Sequence 9 AA;

Query Match 6.8%; Score 8; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 ALAAAYCLS 30
Db 1 ALAAAYCLS 8

RESULT 544
AAU26983
ID AAU26983 standard; peptide; 9 AA.
AC AAU26983;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #267.
XX
KW Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
XX
XX Homo sapiens.
XX OS
XX WO200162776-A1.
PN
PD 30-AUG-2001.
XX
XX 23-FEB-2000; 2000WO-US004655.
XX
XX 23-FEB-2000; 2000WO-US004655.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Sette A, Sidney J, Kast WM, Southwood S;
XX WPI; 2001-582039/65.
XX
XX Composition for treating viral diseases and cancer comprises an
PT immunogenic peptide having an HLA-A2.1 binding motif.
XX
XX Claim 1; Page 68; 85pp; English.
XX
XX Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
CC the invention are capable of specifically binding glycoproteins encoded
CC by HLA alleles and inducing a cytotoxic T cell response against an
CC antigen in a patient expressing HLA-A2.1. This method is useful for the
CC treatment, prevention and diagnosis of pathological states such as viral
CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
CC acuminatum. The peptides are used for treatment of chronic infection and
CC for stimulating the immune system to eliminate virus-infected cells
XX
SQ Sequence 9 AA;

Query Match 6.8%; Score 8; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 ALAAAYCLS 30
Db 1 ALAAAYCLS 8

RESULT 544
AAU26983
ID AAU26983 standard; peptide; 9 AA.
AC AAU26983;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #2543.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
XX Hepatitis C virus.
XX OS
XX WO200121189-A1.
PN
PD 29-MAR-2001.
XX
XX 19-JUL-2000; 2000WO-US019774.
XX
XX 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccine against hepatitis C virus.
XX
XX Disclosure; Page 163; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 9 AA;

Query Match 6.8%; Score 8; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALA 25
Db 2 GGVLAAALA 9

RESULT 546
AAJ01743
ID AAJ01743 standard; peptide; 9 AA.
XX
XX AAJ01743;
XX
XX 02-JUL-2001 (first entry)
XX

AAJ03447
 ID AAJ03447 standard; peptide; 9 AA.
 XX AC AAJ03447;
 XX DT 02-JUL-2001 (first entry)
 XX DE Hepatitis C virus epitope #3438.
 XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 XX KW antiviral.
 XX OS Hepatitis C virus.
 XX PN WO200121189-A1.
 XX PD 29-MAR-2001.
 XX PF 19-JUL-2000; 2000WO-US019774.
 XX PR 19-JUL-1999; 98US-00357737.
 XX PA (EPIM-) EPIMUNE INC.
 XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 XX PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX DR WPI; 2001-308046/32.
 XX PT A new composition useful as a vaccine against hepatitis C virus.
 XX PS Disclosure; Page 178; 214pp; English.
 XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ0010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX Sequence 9 AA;
 Query Match 6.8%; Score 8; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALA 25
 DB 2 GGVLAALA 9
 |||||
 RESULT 550
 ADZ86593
 ID ADZ86593 standard; peptide; 9 AA.
 XX AC ADZ86593;
 XX DT 14-JUL-2005 (first entry)
 XX DE Cytotoxic Hepatitis C virus T-cell epitope, SEQ ID 38.
 XX KW cytotoxic T-lymphocyte; antiinflammatory; virucide; hepatotropic;
 XX KW vaccine; hepatitis C virus infection.
 XX OS Hepatitis C virus.
 XX PN WO2005042698-A2.
 XX PD 12-MAY-2005.
 XX PF 15-OCT-2004; 2004WO-US033942.
 XX PR 23-OCT-2003; 2003US-0513216P.
 XX

(PECO-) PECOS LABS INC.
 Lund O, Lundegaard C, Nielsen M, Worning P, Deans RJ, Buus S;
 Brunak S;
 WPI; 2005-333700/34.
 New T-cell epitope, used as diagnostic tools and as vaccines/composition
 for the treatment and prevention of hepatitis C.
 Claim 1; SEQ ID NO 38; 122pp; English.
 The invention relates to a novel cytotoxic Hepatitis C virus (HCV) T-cell
 epitope, comprising any of the 429 fully defined amino acid sequences, or
 their variations, given in the specification. The invention further
 comprises: a method for predicting peptides that can be used as epitopes
 or as diagnostic tools; a prediction of the neural network combined with
 a prediction or measurement of one of the following: proteasomal cleavage
 sites, MHC binding, presence of sequence or related sequence(s) in patent
 databases, TAP binding, gene or protein expression level, function of the
 protein, or similarity of self proteins; and a vaccine and a diagnostic
 tool, each using a limited number such as at least 1-5, 8, 16, 32, 64,
 128, 256, or 512 of any of the peptides stated above or any of the
 peptides predicted using the method above. The cytotoxic HCV T-cell
 epitopes have antiinflammatory, virucide, and hepatotropic activities.
 The epitopes are useful as vaccines and as diagnostic tools. The vaccines
 and compositions are useful for treating and preventing hepatitis C virus
 infection. This sequence represents one of the 429 cytotoxic Hepatitis C
 virus T-cell epitopes of the invention.
 Sequence 9 AA;
 Query Match 6.8%; Score 8; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CMSADLEV 9
 DB 1 CMSADLEV 8
 |||||
 RESULT 551
 AAW82068
 ID AAW82068 standard; peptide; 10 AA.
 XX AC AAW82068;
 XX DT 18-FEB-1999 (first entry)
 XX DE Fluorogenic protease indicator protease binding peptide #46.
 XX DE Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
 XX KW conformation change.
 XX OS Synthetic.
 XX PN WO9837226-A1.
 XX PD 27-AUG-1998.
 XX PF 20-FEB-1998; 98WO-US003000.
 XX PR 20-FEB-1997; 97US-00802981.
 XX PA (ONCO-) ONCOIMMUNIN INC.
 XX PI Komoriya A, Packard BS;
 XX DR WPI; 1998-467579/40.
 XX PT New fluorogenic compositions - containing 2 fluorophores separated by a
 PT peptide comprising a protease binding site, used for detecting protease
 PT activity in samples.

XX PS Claim 4; Page 77; 90pp; English.

XX CC AAW82023-W82240 are peptides used in the construction of a fluorogenic

CC composition which is used for the detection of protease activity in

CC biological samples. The products can be used for the detection of

CC conformation changes in nucleic acids, oligosaccharides, polysaccharides,

CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,

CC steroids or polymers. In addition, attachment of a hydrophobic group to a

CC molecule can be used to enhance uptake by cells. The composition is

CC composed of P = peptide comprising a protease binding site for the

CC protease, F1, F2 peptides = fluorophores where F1 is attached to the

CC amino terminal amino acid and F2 is attached to the carboxyl terminal

CC amino acid and S1, S2 peptides = when present, are peptide spacers where

CC S1, when present, is attached to the amino terminal acid, and S2, when

CC present, is attached to the carboxyl terminal amino acid

XX SQ Sequence 10 AA;

Query Match 6.8%; Score 8; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67

DB 1 DEMEECSQ 8

RESULT 552

AAJ01891

ID AAJ01891 standard; peptide; 10 AA.

AC AAJ01891;

XX 02-JUL-2001 (first entry)

DE Hepatitis C virus epitope #1882.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;

KW antiviral.

OS Hepatitis C virus.

XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

PI Baker DM, Cellis E, Kubo RT, Grey HM;

XX WPI; 2001-308046/32.

DR 19-JUL-2000; 2000WO-US019774.

XX 19-JUL-1999; 99US-00357737.

XX (EPIM-) EPIMUNE INC.

PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX Baker DM, Cellis E, Kubo RT, Grey HM;

PI WPI; 2001-308046/32.

DR A new composition useful as a vaccines against hepatitis C virus.

XX Disclosure; Page 147; 214pp; English.

XX The present invention describes a composition comprising a prepared

CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.

CC These are derived from HCV HLA-binding motifs. They are useful in

CC vaccines for the prevention and treatment of HCV infection in humans. The

CC present sequence is an epitope used in the disclosure of the invention

XX SQ Sequence 10 AA;

Query Match 6.8%; Score 8; DB 4; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALA 25

DB 3 GGVLAALA 10

RESULT 554

AAJ03750

ID AAJ03750 standard; peptide; 10 AA.

AC AAJ03750;

XX 02-JUL-2001 (first entry)

DE Hepatitis C virus epitope #3741.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;

KW antiviral.

OS Hepatitis C virus.

XX WO200121189-A1.

XX PS Claim 4; Page 77; 90pp; English.

XX CC AAW82023-W82240 are peptides used in the construction of a fluorogenic

CC composition which is used for the detection of protease activity in

CC biological samples. The products can be used for the detection of

CC conformation changes in nucleic acids, oligosaccharides, polysaccharides,

CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,

CC steroids or polymers. In addition, attachment of a hydrophobic group to a

CC molecule can be used to enhance uptake by cells. The composition is

CC composed of P = peptide comprising a protease binding site for the

CC protease, F1, F2 peptides = fluorophores where F1 is attached to the

CC amino terminal amino acid and F2 is attached to the carboxyl terminal

CC amino acid and S1, S2 peptides = when present, are peptide spacers where

CC S1, when present, is attached to the amino terminal acid, and S2, when

CC present, is attached to the carboxyl terminal amino acid

XX SQ Sequence 10 AA;

Query Match 6.8%; Score 8; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67

DB 1 DEMEECSQ 8

RESULT 552

AAJ01891

ID AAJ01891 standard; peptide; 10 AA.

AC AAJ01891;

XX 02-JUL-2001 (first entry)

DE Hepatitis C virus epitope #1882.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;

KW antiviral.

OS Hepatitis C virus.

XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

PI Baker DM, Cellis E, Kubo RT, Grey HM;

XX WPI; 2001-308046/32.

DR 19-JUL-2000; 2000WO-US019774.

XX 19-JUL-1999; 99US-00357737.

XX (EPIM-) EPIMUNE INC.

PA Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX Baker DM, Cellis E, Kubo RT, Grey HM;

PI WPI; 2001-308046/32.

DR A new composition useful as a vaccines against hepatitis C virus.

XX Disclosure; Page 147; 214pp; English.

XX The present invention describes a composition comprising a prepared

CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.

CC These are derived from HCV HLA-binding motifs. They are useful in

CC vaccines for the prevention and treatment of HCV infection in humans. The

CC present sequence is an epitope used in the disclosure of the invention

XX SQ Sequence 10 AA;

Query Match 6.8%; Score 8; DB 4; Length 10;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALA 25

DB 3 GGVLAALA 10

RESULT 554

AAJ03750

ID AAJ03750 standard; peptide; 10 AA.

AC AAJ03750;

XX 02-JUL-2001 (first entry)

DE Hepatitis C virus epitope #3741.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;

KW antiviral.

OS Hepatitis C virus.

XX WO200121189-A1.

XX Protease; indicator; chromophore; H-dimer; fluorescence; absorbance;
KW nuclease; screening; fluorophore; substrate cleavage.
XX Synthetic.
XX WO200261038-A2.
XX 08-AUG-2002.
XX 21-DEC-2001; 2001WO-US049781.
XX 22-DEC-2000; 2000US-00747287.
XX (ONCO-) ONCOIMMUNIN INC.
XX Packard BS, Komoriya A;
XX WPI; 2002-698548/75.
XX Indicator composition comprising polypeptide or nucleic acid backbone
XX joining two same chromophores resulting in quenching of fluorescence
XX of/change in absorbance of chromophores, useful for detecting protease
XX activity.
XX Disclosure; Page 33; 97pp; English.
XX This invention describes a novel indicator composition (referred as homo-
XX doubly labeled compositions) comprising a polypeptide backbone or a
XX nucleic acid backbone joining two chromophores of the same species
XX whereby the chromophores form an H-dimer resulting in quenching of the
XX fluorescence of or a change in the absorbance of the chromophore, a
XX decrease in fluorescence or a change in absorbance indicates that the
XX first molecule and the second molecule are interacting. The indicator is
XX useful for detecting the activity of a protease, where an increase in
XX fluorescence or a change in absorbance indicates that the protease
XX cleaves the polypeptide backbone. The indicator is attached to a solid
XX support inside a mammalian, yeast or insect cell. The composition bears a
XX hydrophobic group such as Fmoc, 9-fluorenesulfonyl group, 1-
XX fluorinecarboxylic group, 9-fluoreneacetyl group, and 9-fluorenone-1-
XX carboxylic group, benzoyloxycarbonyl, Xanthyl (Xan), Trityl (Trt), 4-
XX methyltrityl (Mt), 4-methoxytrityl (Mmt), 4-methoxy-2,3,6-trimethyl-
XX benzenesulfonyl (Mtr), Mesitylene-2-sulfonyl (Mts), 4,4'-
XX dimethoxybenzhydryl (Mbh), etc. The method described in the invention is
XX useful for detecting protease or nuclease activity (or the presence of
XX nucleic acid) in histological section, cells in culture, (e.g., seeded or
XX cultured adherent cells), a biological sample such as tissue, biopsy,
XX lymph, embryo, or whole animal, or cell suspension derived from a
XX biological sample such as tissue, blood, urine, saliva, lymph, or biopsy.
XX The indicator composition is also useful for screening a test agent for
XX the ability to modulate a protease (or a nuclease, lipase, etc.). The
XX indicator reagents allow rapid determination of protease activity in a
XX matter of minutes in a single-step procedure. The fluorescent indicators
XX both absorb and emit in the visible range (400-800 nm). These signals are
XX therefore not readily quenched by, nor is activation of the fluorophores,
XX that is, absorption of light, interfered with by background molecules;
XX therefore they are easily detected in biological samples. The fluorogenic
XX protease indicators utilize high efficiency fluorophores and are able to
XX achieve a high degree of quenching while providing a strong signal when
XX the quench is released by cleavage of the peptide substrate. The high
XX signal allows detection of very low levels of protease activity. Thus the
XX fluorogenic protease indicators are particularly well suited for in situ
XX detection of protease activity. ABU60357-ABU60477 represent peptides use
XX to illustrate the method described in the disclosure of the invention
XX Sequence 10 AA;
XX
XX Query Match 6.8%; Score 8; DB 5; Length 10;
XX Best Local Similarity 100.0%; Pred. No. 2.7;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 60 DEMEECSQ 67
XX |||||

Db 1 DEMEECSQ 8
RESULT 558
AAW71278
ID AAW71278 standard; peptide; 11 AA.
XX
XX AAW71278;
AC 17-NOV-1998 (first entry)
DT
XX
XX Cleavable substrate sequence.
DE
XX
XX HCV NS3 protease; NS4A cofactor protein; HCV therapy; fusion protein.
KW
XX
XX Synthetic.
OS
XX
XX WO9837180-A2.
PN
XX
XX 27-AUG-1998.
PD
XX
XX 20-FEB-1998; 98WO-US003367.
PF
XX
XX 22-FEB-1997; 97US-00804266.
PR
XX
XX (ABBO) ABBOTT LAB.
PA
XX
XX Chen C, Molla A, Tripathi RL;
PI
XX
XX WPI; 1998-467551/40.
DR
XX
XX New hepatitis C virus fusion proteins - comprises NS3 protease and NS4A
PT co-factor, used in assays for screening for compounds for use in HCV
PT therapy.
XX
XX Disclosure; Page 6; 31pp; English.
XX
XX AAW71273-83 represent cleavable substrate peptide sequences. The
CC specification describes a fusion protein derived from the Hepatitis C
CC virus (HCV) NS3 protease and NS4A cofactor proteins (NS3/4A). A non-
CC autocleavable fusion protein of HCV NS3 protease and HCV NS4A cofactor
CC protein is produced upon expression, which is biologically active. The
CC products can be used to obtain drugs, which can inhibit NS3 protease
CC activity for use in HCV therapy. They can also be used to design
CC compounds which interact with or inhibit the NS3/4A fusion proteins
XX
XX Sequence 11 AA;
XX
XX Query Match 6.8%; Score 8; DB 2; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 2.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 60 DEMEECSQ 67
XX |||||
Db 1 DEMEECSQ 8
RESULT 559
AAJ00663
ID AAJ00663 standard; peptide; 11 AA.
XX
XX AAJ00663;
AC
XX
XX 02-JUL-2001 (first entry)
DT
XX
XX Hepatitis C virus epitope #654.
DE
XX
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
XX Hepatitis C virus.
OS
XX
XX WO200121189-A1.
PN

XX 29-MAR-2001.
XX
XX
XX 19-JUL-2000; 2000WO-US019774.
XX
XX 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
XX
XX Disclosure; Page 116; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 11 AA;
SQ

Query Match 6.8%; Score 8; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALA 25
Db 4 GGVLAALA 11

RESULT 560
AAJ02576
ID AAJ02576 standard; peptide; 11 AA.
XX
XX AAJ02576;
XX
XX 02-JUL-2001 (first entry)
XX
XX Hepatitis C virus epitope #2567.
XX
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX Hepatitis C virus.
XX
XX WO200121189-A1.
XX
XX 29-MAR-2001.
XX
XX 19-JUL-2000; 2000WO-US019774.
XX
XX 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
XX
XX Disclosure; Page 163; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The

CC present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 11 AA;
XX
XX Query Match 6.8%; Score 8; DB 4; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 2.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALA 25
Db 4 GGVLAALA 11

RESULT 561
AAJ02019
ID AAJ02019 standard; peptide; 11 AA.
XX
XX AAJ02019;
XX
XX 02-JUL-2001 (first entry)
XX
XX Hepatitis C virus epitope #2010.
XX
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX
XX Hepatitis C virus.
XX
XX WO200121189-A1.
XX
XX 29-MAR-2001.
XX
XX 19-JUL-2000; 2000WO-US019774.
XX
XX 19-JUL-1999; 99US-00357737.
XX
XX (EPIM-) EPIMMUNE INC.
XX
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX
XX A new composition useful as a vaccines against hepatitis C virus.
XX
XX Disclosure; Page 150; 214pp; English.
XX
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX
XX Sequence 11 AA;
XX
XX Query Match 6.8%; Score 8; DB 4; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 2.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALA 25
Db 4 GGVLAALA 11

RESULT 562
AEA39175
ID AEA39175 standard; peptide; 13 AA.
XX
XX AEA39175;
XX
XX 25-AUG-2005 (first entry)
XX
XX Hepatitis C virus specific protease scission site SEQ ID NO 73.

CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC C virus protein.
XX
SQ Sequence 18 AA;

Query Match 6.8%; Score 8; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
Db 5 DEMEECSQ 12
|||||

RESULT 566
ADV22997
ID ADV22997 standard; peptide; 18 AA.
XX
AC ADV22997;
XX
DT 10-MAR-2005 (first entry)
XX
DE HCV H77 immunogenic peptide #238.
XX
KW Vaccine; viricide; antigen; autoimmune disease; infection;
KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
KW hodgekin's lymphoma.
XX
OS Hepatitis C virus strain H77.
XX
PN WO2004108753-A1.
XX
PD 16-DEC-2004.
XX
PP 10-JUN-2004; 2004WO-AU000775.
XX
PR 10-JUN-2003; 2003AU-00902875.
PR 25-MAR-2004; 2004AU-00901589.
XX
PA (UYME) UNIV MELBOURNE.
XX
PI Kent SJ;
XX
DR WPI; 2005-031657/03.
XX
PT Use of at least one set of peptides in the preparation of a medicament
PT for modulating an immune response, and for treating cancer or yeast,
PT viral, bacterial, protozoal and mycoplasma infections.
XX
PS Disclosure; SEQ ID NO 1417; 645pp; English.
XX
CC The invention relates to the use of at least one set of peptides in the
CC preparation of a medicament for modulating an immune response, where
CC individual peptides of a respective set comprise different portions of an
CC amino acid sequence corresponding to a single polypeptide of interest and
CC display partial sequence identity or similarity to at least one other
CC peptide of the same set of peptides (i.e. they are overlapping). Also
CC included are an antigen-presenting cell which has been contacted with the
CC peptides above and thus presents the peptides, a population of such
CC antigen-presenting cells, a process for producing antigen-presenting
CC cells for modulating an immune response to a polypeptide of interest, a
CC method for producing antigen-specific lymphocytes, a composition
CC comprising at least one set of the peptides (and a carrier and/or
CC diluent), a method for modulating an immune response to a polypeptide of
CC interest comprising administering to a patient in need at least one set
CC of the peptides, a method for treatment and/or prophylaxis of a disease
CC or condition associated with the presence of a polypeptide of interest
CC and a composition of matter for modulating an immune response in a
CC subject to a target antigen. The polypeptide of interest is also a
CC disease- or condition-associated polypeptide that is a polypeptide
CC produced by a pathogenic organism or a cancer, and produced by a

CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
CC protozoans and mycoplasmas. The disease- or condition-associated
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC of aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC C virus protein.
XX
SQ Sequence 18 AA;

Query Match 6.8%; Score 8; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db 10 CMSADLEV 17
|||||

RESULT 567
AAW82168
ID AAW82168 standard; peptide; 19 AA.
XX
AC AAW82168;
XX
DT 18-FEB-1999 (first entry)
XX
DE Fluorogenic protease indicator NS3NS4A/4B peptide #1.
XX
KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
KW conformation change.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 3 /label= Alb
FT FT /note= "alpha-aminoisobutyric acid, labelled as amino
FT FT acid B in the specification"
FT Modified-site 4
FT FT /note= "epsilon-aminocaproic acid, labelled as amino acid
FT FT J in the specification"
XX
XX WO9837226-A1.
XX
XX 27-AUG-1998.
XX
XX 20-FEB-1998; 98WO-US003000.
XX
XX 20-FEB-1997; 97US-00802981.
XX
XX (ONCO-) ONCOIMMUNIN INC.
XX
XX Komoriya A, Packard BS;
XX WPI; 1998-467579/40.
XX
XX New fluorogenic compositions - containing 2 fluorophores separated by a
XX peptide comprising a protease binding site, used for detecting protease
XX activity in samples.
XX
XX Disclosure; Page 27; 90pp; English.
XX
XX AAW82023-W82240 are peptides used in the construction of a fluorogenic

CC composition which is used for the detection of protease activity in
CC biological samples. The products can be used for the detection of
CC conformation changes in nucleic acids, oligosaccharides, polysaccharides,
CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,
CC steroids or polymers. In addition, attachment of a hydrophobic group to a
CC molecule can be used to enhance uptake by cells. The composition is
CC composed of P = peptide comprising a protease binding site for the
CC protease, F1, F2 peptides = fluorophores where F1 is attached to the
CC amino terminal amino acid and F2 is attached to the carboxyl terminal
CC amino acid and S1, S2 peptides = when present, are peptide spacers where
CC S1, when present, is attached to the amino terminal acid, and S2, when
CC present, is attached to the carboxyl terminal amino acid
XX
SQ Sequence 19 AA;

Query Match 6.8%; Score 8; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 6 DEMEECSQ 13

RESULT 568

AAG73167
ID AAG73167 standard; peptide; 19 AA.

XX AC AAG73167;

XX DT 14-AUG-2001 (first entry)

XX DE Protease binding site #101.

XX KW Protease detection; peptide cleavage; enzyme activity; fluorogenic;
XX KW viral infection; cancer metastasis; emphysema; arthritis; thrombosis;
XX KW haemophilia.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 3

XX FT /label= Aib

XX FT /note= "2-aminoisobutyric acid"

XX FN WO200118238-A1.

XX PD 15-MAR-2001.

XX PF 11-SEP-2000; 2000WO-US024882.

XX PR 10-SEP-1999; 99US-00394019.

XX PA (ONCO-) ONCOIMMUNIN INC.

XX PI Komoriya A, Packard BS;

XX DR WPI; 2001-389573/41.

XX PT New fluorogenic compositions whose fluorescence level increases in the
XX PT presence of active proteases, useful for detecting and localizing
XX PT protease activity in biological samples, particularly in frozen tissue
XX PT samples.

XX PS Disclosure; Page 26; 86pp; English.

XX CC The present invention describes fluorogenic compositions which can be
XX CC used for the detection of protease activity. This can be useful as an
XX CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
XX CC thrombosis and arthritis. The fluorogenic compositions comprise a
XX CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
XX CC peptide is cleaved by a protease and the fluorophores can then be
XX CC detected. The present sequence is one of the peptides described in the

CC exemplification of the invention

XX SQ Sequence 19 AA;

Query Match 6.8%; Score 8; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67

| | | | | | | |

Db 6 DEMEECSQ 13

RESULT 569

ADN88407

ID ADN88407 standard; peptide; 19 AA.

XX AC ADN88407;

XX DT 12-AUG-2004 (first entry)

XX DE Fluorogenic protease indicator peptide #103.

XX KW fluorogenic; protease detection; protease inhibitor.

XX OS Synthetic.

XX FN US2004096926-A1.

XX PD 20-MAY-2004.

XX PF 04-JUN-2001; 2001US-00874350.

XX PR 20-FEB-1997; 97US-00802981.

XX PR 20-FEB-1998; 98WO-US003000.

XX PR 10-SEP-1999; 99US-00394019.

XX PR 11-SEP-2000; 2000WO-US024882.

XX PA (ONCO-) ONCOIMMUNIN INC.

XX PI Packard BS, Komoriya A;

XX DR WPI; 2004-399235/37.

XX PT Fluorogenic composition useful for detecting protease activity and test
XX PT substance modulating protease activity.

XX PS Disclosure; SEQ ID NO 103; 114pp; English.

XX CC The invention relates to a fluorogenic composition (I) for detecting the
XX CC activity of a protease. (I) is useful for detecting the activity of a
XX CC protease, which involves contacting the protease with (I), where the
XX CC activity of protease is detected in a histological section, cell culture
XX CC or tissue section. The cell suspension is derived from the biological
XX CC sample chosen from tissue, blood, urine, saliva, lymph or biopsy. The
XX CC protease activity is detected by fluorescence microscopy, fluorescence
XX CC microplate reader, absorption microplate reader, flow cytometry,
XX CC fluorometry, absorption spectroscopy or confocal fluorescent microplate
XX CC reader. (I) is useful for delivering a molecule into a cell, and for
XX CC screening a test agent for the ability to modulate the activity of the
XX CC protease. (I) is useful for detection and localisation of protease
XX CC activity in biological samples. (I) also acts as a protease inhibitor,
XX CC thus useful as protease inhibitors. (I) enables detection of the protease
XX CC activity, and provides a high intensity fluorescent signal at a visible
XX CC wavelength when they are digested by a protease. The present sequence
XX CC represents a fluorogenic protease indicator peptide of the invention.

XX SQ Sequence 19 AA;

Query Match 6.8%; Score 8; DB 8; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 60 DEMEECSQ 67
DB 6 DEMEECSQ 13

RESULT 570
AAR62943
ID AAR62943 standard; peptide; 20 AA.
XX AC AAR62943;
XX AC 25-MAR-2003 (revised)
XX DT 04-AUG-1995 (first entry)
XX XX Hepatitis C virus NS4 protein synthetic peptide.
XX XX Synthetic peptide; solid phase immunoassay.
XX XX Synthetic.
XX PN WO9426932-A1.
XX XX 24-NOV-1994.
XX XX 13-MAY-1994; 94WO-US005407.
XX XX 13-MAY-1993; 93US-00061694.
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX XX Fields HA, Khudyakov YE;
XX XX WPI; 1995-006819/01.
XX XX Solid phase immunoassay using oligo:nucleotide as label - also new
XX PT conjugates of oligo:nucleotide coupled to antigenic peptide, partic. for
XX PT diagnosing hepatitis C or E virus infection.
XX XX Example; Page 18; 34pp; English.
XX XX AAR62941 and AAR62942 are examples of synthetic immunoreactive peptides.
XX CC They are used in a method for detecting an antigen in a subject. The
XX CC method involves binding the antigen to a solid support and then reacting
XX CC it with an immunoreactive ligand (L) bound to an oligo; removing any
XX CC unreacted L, and then detecting the presence of the oligo. A similar
XX CC method can be used to detect Abs, in which case the ligand is an oligo-
XX CC labelled Ag. The use of an amplifiable oligo as the label allows Ag or Ab
XX CC to be detected at very low levels. In the example, a synthetic peptide
XX CC from the NS4 protein of the hepatitis C virus with structure (AAR62943)
XX CC is biotinylated using a commercially available kit. A biotinylated oligo
XX CC with the structure 5'-biotinylated-AAQ75033-3' was prep'd. This oligo is
XX CC composed of sequences of two PCR primers sepd. by a short additional
XX CC sequence. The shorter the region to be amplified the better the
XX CC efficiency of amplification obtd. The biotinylated oligo is pre-
XX CC incubated with streptavidin. Then this complex linked by biotin-
XX CC streptavidin binding. This INAA complex is then used in place of
XX CC chemically prep'd. oligo-peptide conjugates mentioned above. (Updated on
XX CC 25-MAR-2003 to correct PN field.)
XX XX Sequence 20 AA;

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 13 DEMEECSQ 20

RESULT 571
AAR10132
ID AAR10132 standard; peptide; 20 AA.
XX AC AAR10132;
XX AC 25-MAR-2003 (revised)
XX DT 02-OCT-1997 (first entry)
XX XX Hepatitis C virus peptide antigen IX.
XX XX Antibody; HCV; immunoassay; vaccine; mimic.
XX XX Synthetic.
XX OS
XX PH Key
XX FT Modified-site 1
XX FT /note= "H or a linker arm by which the peptide can be
XX FT attached to a carrier or solid phase comprising at
XX FT least one amino acid and as many as 60, most frequently 1-20
XX FT amino acids, such as Cys, Lys, Tyr, Glu or Asp, or
XX FT chemical groups such as biotin or thioglycolic acid; can
XX FT be modified by acetylation"
XX FT Modified-site 20
XX FT /note= "A bond or a linker arm by which the peptide can
XX FT be attached to a carrier or solid phase comprising at
XX FT least one amino acid and as many as 60 amino acids, most
XX FT frequently 1-10 amino acids, such as Cys, Lys, Tyr, Asp,
XX FT or chemical groups such as biotin or thioglycolic acid;
XX FT and attached on to that is NH2, OH or a linkage involving
XX FT either of these two groups"
XX XX EP754704-A2.
XX PN
XX XX 22-JAN-1997.
XX PD
XX XX 14-DEC-1990; 96EP-00201157.
XX PF
XX XX 14-DEC-1990; 90EP-00124241.
XX PR
XX XX (INNO-) INNOGENETICS NV.
XX PA
XX XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX PI WPI; 1997-089256/09.
XX DR Hepatitis C virus peptide mimics - for use in immunoassays, vaccines,
XX FT etc.
XX XX Claim 1; Page 34; 65pp; English.
XX PS The present sequence represents a novel synthetic Hepatitis C virus (HCV)
XX CC antigen IX for the detection of antibodies. The peptide contains
XX CC modifications at the N- and C-terminal (see features table) with the
XX CC condition that if the modification represents an amino acid(s), that they
XX CC are different from any naturally occurring HCV flanking regions. The
XX CC peptide represents an HCV peptide mimic and may be used as an immunoassay
XX CC reagent for detecting antibodies to HCV; for incorporation into vaccines
XX CC against HCV; and for raising antibodies against HCV. (Updated on 25-MAR-
XX CC 2003 to correct PF field.)
XX XX Sequence 20 AA;

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 13 DEMEECSQ 20

RESULT 572
AAR10133
ID AAR10133 standard; peptide; 20 AA.
XX AC AAR10133;
XX XX

```

AC AAW10133;
 XX 25-MAR-2003 (revised)
 DT 02-OCT-1997 (first entry)
 XX
 XX Hepatitis C virus peptide antigen X.
 XX Antibody; HCV; immunoassay; vaccine; mimic.
 XX Synthetic.
 XX OS
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "H or a linker arm by which the peptide can be
 FT attached to a carrier or solid phase comprising at least
 FT one amino acid and as many as 60, most frequently 1-20
 FT amino acids, such as Cys, Lys, Tyr, Glu or Asp, or
 FT chemical groups such as biotin or thiolglycolic acid; can
 FT be modified by acetylation"
 FT 20
 FT Modified-site
 FT /note= "A bond or a linker arm by which the peptide can
 FT be attached to a carrier or solid phase comprising at
 FT least one amino acid and as many as 60 amino acids, most
 FT frequently 1-10 amino acids, such as Cys, Lys, Tyr, Asp,
 FT or chemical groups such as biotin or thiolglycolic acid;
 FT and attached on to that is NH2, OH or a linkage involving
 FT either of these two groups"
 XX
 XX EP754704-A2.
 PN
 XX 22-JAN-1997.
 PD
 XX 14-DEC-1990; 96EP-00201157.
 XX
 XX 14-DEC-1990; 90EP-00124241.
 XX
 XX (INNO-) INNOGENETICS NV.
 XX
 XX Deleys RJ, Follet D, Maertens G, Van Heuverswyn H;
 PI WPI; 1997-089256/09.
 XX
 XX Hepatitis C virus peptide mimics - for use in immunoassays, vaccines,
 FT etc.
 FT
 PS Claim 1; Page 34; 65pp; English.
 XX
 XX The present sequence represents a novel synthetic Hepatitis C virus (HCV)
 CC antigen X for the detection of antibodies. The peptide contains
 CC modifications at the N- and C-terminal (see features table) with the
 CC condition that if the modification represents an amino acid(s), that they
 CC are different from any naturally occurring HCV flanking regions. The
 CC peptide represents an HCV peptide mimic and may be used as an immunoassay
 CC reagent for detecting antibodies to HCV; for incorporation into vaccines
 CC against HCV; and for raising antibodies against HCV. (Updated on 25-MAR-
 CC 2003 to correct PF field.)
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 6.8%; Score 8; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMERCSQ 67
 |||||
 Db 1 DEMERCSQ 8
 RESULT 573
 AAW47150
 ID AAW47150 standard; peptide; 20 AA.
 XX
 AC AAW47150;

XX 26-MAY-1998 (first entry)
 DT
 XX Hepatitis C virus (HCV) NS3/4A substrate fragment.
 DE
 XX Hepatitis C virus; HCV; HCV NS3 protease; inhibitor; bivalent;
 KW monovalent; HCV NS4A co-factor; HCV infection; NS3/4A substrate.
 XX
 XX Hepatitis C virus.
 OS
 XX Key Location/Qualifiers
 FH Cleavage-site 10.11
 FT WO9743310-A1.
 XX
 XX 20-NOV-1997.
 PD
 XX 08-MAY-1997; 97WO-US007632.
 PF
 XX 10-MAY-1996; 96US-00644544.
 PR
 XX (SCHE) SCHERING CORP.
 XX
 XX Zhang R, Mui PW, Weber PC;
 PI WPI; 1998-008797/01.
 XX
 XX New inhibitors of hepatitis C protease NS3 - contain at least one of NS3
 PT substrate sequence and NS4A co-factor polypeptide, for treatment of
 FT hepatitis C infection.
 FT
 XX Disclosure; Page 6; 59pp; English.
 PS
 XX This is a NS3/4A substrate fragment of hepatitis C virus (HCV) containing
 CC a NS3 protease cleavage site. This can be used to construct inhibitors of
 CC HCV NS3 protease and in assays for determining the inhibitory activity of
 CC the novel bivalent and monovalent HCV NS3 protease inhibitors. A novel
 CC bivalent inhibitor of a NS3 protease comprises a first peptide that is a
 CC subsequence, mutated subsequence or a mutated full-length sequence of the
 CC NS3 substrate linked to a second peptide that is a subsequence of the HCV
 CC NS4A polypeptide. These bivalent inhibitors and other monovalent
 CC inhibitors of an HCV protease comprising a subsequence, mutated
 CC subsequence or a mutated full-length sequence of the substrate of HCV NS3
 CC protease or a subsequence, mutated subsequence or mutated full-length
 CC sequence of NS4A are used to treat HCV infection. They act by inhibiting
 CC the interaction between NS3 and at least one of its substrates and the
 CC NS4A co-factor. Compared with inhibitors that target only one component,
 CC the bivalent inhibitors may have higher binding affinity and better
 CC discrimination against similar host cell enzymes, i.e. reduced toxicity.
 CC The peptide inhibitors can be assessed for their inhibitory activity by a
 CC scintillation proximity assay using NS3, NS4 and peptide substrates 4B/5A
 CC or 5A/5B. The inhibitors are made by usual methods of solid phase
 CC synthesis and can be administered orally or by injection or by
 CC transdermal diffusion, optionally conjugated to a carrier protein
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 6.8%; Score 8; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CMSADLEV 9
 |||||
 Db 1 CMSADLEV 8
 RESULT 574
 AAW47151
 ID AAW47151 standard; peptide; 20 AA.
 XX
 AC AAW47151;
 XX
 XX 26-MAY-1998 (first entry)
 DT

XX DE Hepatitis C virus (HCV) NS4A/4B substrate fragment.
XX KW Hepatitis C virus; HCV; HCV NS3 protease; inhibitor; bivalent;
XX KW monovalent; HCV NS4A co-factor; HCV infection; NS4A/4B substrate.
XX OS Hepatitis C virus.
XX PH Key Location/Qualifiers
XX FT Cleavage-site 10. .11
XX PN WO9743310-A1.
XX XX 20-NOV-1997.
XX PF 08-MAY-1997; 97WO-US007632.
XX PR 10-MAY-1996; 96US-00644544.
XX PA (SCHE) SCHERING CORP.
XX PI Zhang R, Mui PW, Weber PC;
XX PT WPI; 1998-008797/01.
XX DR
XX PT New inhibitors of hepatitis C protease NS3 - contain at least one of NS3
XX PT substrate sequence and NS4A co:factor polypeptide, for treatment of
XX PT hepatitis C infection.
XX PS Disclosure; Page 7; 59pp; English.
XX CC This NS4A/4B substrate fragment of hepatitis C virus (HCV) contains a NS3
XX CC protease cleavage site. This can be used to construct inhibitors of HCV
XX CC NS3 protease and in assays for determining the inhibitory activity of the
XX CC novel bivalent and monovalent HCV NS3 protease inhibitors. A novel
XX CC bivalent inhibitor of a NS3 protease comprises a first peptide that is a
XX CC subsequence, mutated subsequence or a mutated full-length sequence of the
XX CC NS3 substrate linked to a second peptide that is a subsequence of the HCV
XX CC NS4A polypeptide. These bivalent inhibitors and other monovalent
XX CC inhibitors of an HCV protease comprising a subsequence, mutated
XX CC subsequence or a mutated full-length sequence of the substrate of HCV NS3
XX CC protease or a subsequence, mutated subsequence or mutated full-length
XX CC sequence of NS4A are used to treat HCV infection. They act by inhibiting
XX CC the interaction between NS3 and at least one of its substrates and the
XX CC NS4A co-factor. Compared with inhibitors that target only one component,
XX CC the bivalent inhibitors may have higher binding affinity and better
XX CC discrimination against similar host cell enzymes, i.e. reduced toxicity.
XX CC The peptide inhibitors can be assessed for their inhibitory activity by a
XX CC scintillation proximity assay using NS3, NS4 and peptide substrates 4B/5A
XX CC or 5A/5B. The inhibitors are made by usual methods of solid phase
XX CC synthesis and can be administered orally or by injection or by
XX CC transdermal diffusion, optionally conjugated to a carrier protein
XX SQ Sequence 20 AA;
Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 5 DEMEECSQ 12
RESULT 575
AAY57210 6.8%; Score 8; DB 2; Length 20;
ID AAY57210 standard; peptide; 20 AA.
XX AC AAY57210;
XX DT 29-FEB-2000 (first entry)
XX DE NS4A/4B cleavage site.
RESULT 576
AAY57209 6.8%; Score 8; DB 3; Length 20;
ID AAY57209 standard; peptide; 20 AA.
XX AC AAY57209;
XX DT 29-FEB-2000 (first entry)
XX DE NS3/4A cleavage site.
Hepatitis C virus; HCV; NS3 protease; bivalent inhibitor; linker;
NS4A polypeptide; monovalent inhibitor.
Hepatitis C virus.
US5990276-A.
23-NOV-1999.
09-MAY-1997; 97US-00853623.
10-MAY-1996; 96US-0017470P.
(SCHE) SCHERING CORP.
Zhang R, Mui PW, Weber PC;
WPI; 2000-037868/03.
Bivalent and monovalent inhibitors of hepatitis C virus NS3 protease.
Disclosure; Col 3-4; 27pp; English.
The invention provides bivalent inhibitors of hepatitis C virus (HCV) NS3
protease. The bivalent inhibitor comprises: (a) a first peptide
consisting of a subsequence, a mutated subsequence or a mutated full-
length sequence of a substrate of the HCV NS3 protease which is not
cleaved by the protease; (b) a second peptide consisting of a subsequence
of a HCV NS4A polypeptide (sequences AY57195-201); (c) a linker
comprising a chemical entity capable of forming a bond with the first
peptide and the second peptide and is equivalent in length to a carbon
chain having 7-14 carbon atoms. Monovalent inhibitors of the HCV NS3
protease inhibit either the interaction of a substrate or the cofactor
NS4A with the NS3 protease, and the bivalent inhibitor inhibits the
interaction of the NS3 protease with both cofactor NS4A and a substrate
of the NS3 protease. The mono- and bivalent inhibitors are useful for
treating an individual infected with the HCV. The bivalent enzyme
inhibitors provide a higher binding affinity (potency), as well as
enhanced specificity against similar cellular host enzymes for reduced
toxicity effects. The present sequence represents the NS4A/4B cleavage
site
SQ Sequence 20 AA;
Query Match 6.8%; Score 8; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 5 DEMEECSQ 12
RESULT 576
AAY57209 6.8%; Score 8; DB 3; Length 20;
ID AAY57209 standard; peptide; 20 AA.
XX AC AAY57209;
XX DT 29-FEB-2000 (first entry)
XX DE NS3/4A cleavage site.
Hepatitis C virus; HCV; NS3 protease; bivalent inhibitor; linker;
NS4A polypeptide; monovalent inhibitor.
Hepatitis C virus.
US5990276-A.
23-NOV-1999.
XX

PF 09-MAY-1997; 97US-00853623.
 PR 10-MAY-1996; 96US-0017470P.
 XX (SCHE) SCHERING CORP.
 XX Zhang R, Mui PW, Weber PC;
 XX WPI; 2000-037868/03.
 XX Bivalent and monovalent inhibitors of hepatitis C virus NS3 protease.
 XX Disclosure; Col 3-4; 27pp; English.
 XX The invention provides bivalent inhibitors of hepatitis C virus (HCV) NS3
 CC protease. The bivalent inhibitor comprises: (a) a first peptide
 CC consisting of a subsequence, a mutated subsequence or a mutated full-
 CC length sequence of a substrate of the HCV NS3 protease which is not
 CC cleaved by the protease; (b) a second peptide consisting of a subsequence
 CC of a HCV NS4A polypeptide (sequences AY57195-201); (c) a linker
 CC comprising a chemical entity capable of forming a bond with the first
 CC peptide and the second peptide and is equivalent in length to a carbon
 CC chain having 7-14 carbon atoms. Monovalent inhibitors of the HCV NS3
 CC protease inhibit either the interaction of a substrate or the cofactor
 CC NS4A with the NS3 protease, and the bivalent inhibitor inhibits the
 CC interaction of the NS3 protease with both cofactor NS4A and a substrate
 CC of the NS3 protease. The mono- and bivalent inhibitors are useful for
 CC treating an individual infected with the HCV. The bivalent enzyme
 CC inhibitors provide a higher binding affinity (potency), as well as
 CC enhanced specificity against similar cellular host enzymes for reduced
 CC toxicity effects. The present sequence represents the NS3/4A cleavage
 CC site
 XX
 SQ Sequence 20 AA;
 Query Match 6.8%; Score 8; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 CMSADLEV 9
 Db | | | | | | | |
 1 CMSADLEV 8
 RESULT 577
 AAY93535
 ID AAY93535 standard; peptide; 20 AA.
 XX
 AC AAY93535;
 XX
 DT 25-SEP-2000 (first entry)
 XX
 DE Amino acid sequence of a HCV protease cleavage site.
 XX
 KW Protease cleavage site; protein transduction system;
 KW protein transduction domain; cytotoxic domain; pathogen infection;
 KW retroviral infection; plasmoidal infection; cancer; prostate cancer.
 XX
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT Cleavage-site 10.11
 FT /note= "cleavage site from a NS3-NS4A protein"
 XX
 XX WO200034308-A2.
 XX
 PD 15-JUN-2000.
 XX
 XX 10-DEC-1999; 99WO-US029289.
 XX
 XX 10-DEC-1998; 98US-0111701P.
 XX
 XX (UNIW) UNIV WASHINGTON.

XX Dowdy SF;
 PI WPI; 2000-431269/37.
 XX
 DR Protein transduction system for treating cancer and pathogenic infections
 PT has a fusion protein comprising a protein transduction domain covalently
 PT linked to a cytotoxic domain.
 XX
 PS Disclosure; Page 46; 127pp; English.
 XX
 CC The present sequence represents a cleavage site sequence from HCV
 CC protein. The present protease cleavage site is used in the protein
 CC transduction system of the invention. The specification describes a
 CC protein transduction system, which comprises a fusion protein. This
 CC fusion protein has a covalently linked protein transduction domain and
 CC cytotoxic domain. The system is useful for treating pathogen infection in
 CC mammals, infections such as those caused by CMV, HSV-1, HCV, KSHV, yellow
 CC fever virus, flavivirus or rhinovirus, retroviral infections such as HIV-
 CC 1, HIV-2, HTLV-3 and/or LAV, plasmoidal infections associated with
 CC P.faciapum, P.vivax, P.ovale, P.malariae. It is also useful for treating
 CC cancer, especially prostate cancer
 XX
 SQ Sequence 20 AA;
 Query Match 6.8%; Score 8; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 CMSADLEV 9
 Db | | | | | | | |
 1 CMSADLEV 8
 RESULT 578
 AAB29428
 ID AAB29428 standard; peptide; 20 AA.
 XX
 AC AAB29428;
 XX
 DT 09-FEB-2001 (first entry)
 XX
 DE HCV NS3-NS4A cleavage site.
 XX
 KW Protein transduction domain; fusion molecule; therapeutic agent;
 KW drug targeting; drug discovery; cell transduction; bioavailability;
 KW vaccine; nervous system disorder; Alzheimer's disease;
 KW Parkinson's disease; Huntington's disease; pre-senile dementia; epilepsy;
 KW seizure; compulsive behaviour; meningitis; encephalitis; ischaemia;
 KW spongiform encephalopathy; dyslexia; age-related memory loss;
 KW Lou Gehring's disease; viral infection; HIV; bacterial infection;
 KW protease cleavage site.
 XX
 OS Hepatitis C virus.
 XX
 FN WO200062067-A1.
 XX
 XX 19-OCT-2000.
 XX
 XX 28-FEB-2000; 2000WO-US005097.
 XX
 XX 28-FEB-1999; 99US-0122757P.
 PR 29-AUG-1999; 99US-0151291P.
 XX
 XX (UNIW) UNIV WASHINGTON.
 PA
 XX Dowdy SF;
 XX WPI; 2000-647439/62.
 XX
 FT Fusion molecules comprising protein transduction domains and therapeutic
 PT agents, useful for treating e.g. Alzheimer's and Parkinson's diseases,
 PT dementia and epilepsy.

XX PS Disclosure; Page 63; 19pp; English.

XX CC The invention relates to a novel fusion molecule comprising at least one protein transduction domain (PTD) and at least one linked molecule, where the linked molecule has therapeutic or prophylactic activity against a medical condition. The invention also relates to methods of drug discovery in which the test compound is linked to a suitable transducing protein and introduced to a cell; a method of killing resistant microorganisms using a suitable fusion molecule; a mammal comprising a covalently linked fusion molecule; and a mammal adapted for experimental use in which at least one transduction molecule has been transduced into essentially all the cells of the mammal. The fusion molecule is used to deliver a therapeutic agent to a mammal, especially a human. The linked molecule may be a vaccine, an anti-infective drug, a cardiovascular drug, an antitumour drug, an analgesic, an antiinflammatory, a diagnostic marker or a drug for the treatment or prevention of a central or peripheral nervous system disorder. The central nervous system (CNS) disorder is especially Alzheimer's disease, Parkinson's disease, Huntington's disease, and also includes pre-senile dementia, epilepsy and seizures, compulsive behaviour, meningitis (including viral and bacterial meningitis), encephalitis, ischaemia, scrapie (or related spongiform encephalopathies), dyslexia, age-related memory loss or Lou Gehring's disease. Fusion molecules can also be used to kill virally infected cells, especially those infected with HIV. The vaccines are used to treat or prevent bacterial or viral infections. The methods are a highly effective means for transducing a molecule into an entire mammal or into specific cells, tissues, organs and systems within it. They also overcome bioavailability problems that are associated with many therapeutic agents (e.g., large molecular size, hydrophobicity, hydrophilicity, biological resistance), by providing efficient transduction of the target cell. The present sequence represents a protease cleavage site which may be used in the fusion molecules of the invention

XX SQ Sequence 20 AA;

Query Match 6.8%; Score 8; DB 3; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
 Db 1 CMSADLEV 8
 |||||

RESULT 579
 AAR41108
 ID AAR41108 standard; peptide; 22 AA.
 AC AAR41108;
 XX 25-MAR-2003 (revised)
 DT 22-MAR-1994 (first entry)
 XX HCV peptide HCV3.
 DE Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 1 /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B= biotin; X= biotinylation cpd. incorporated during synthesis; Y= bond or linking gp(s). which minimises steric hindrance, where Y is not a bond it is pref. 1-10 residues of (same or different) glycine, beta-alanine, 4-aminobutyric acid, 5-aminovaleric acid or 6-aminohexanoic acid; parentthesis around B and X indicate opt. presence at the specified positions but B or X must be present in at least one of the positions shown, B interacts with the

FT peptide to give a cpd. with greater diagnostic sensitivity; A (optional)= one or more amino acids, NH2 or gp. which modifies the N-terminus; Z= one or more amino acids, OH, NH2, or a linkage involving either of these 2 gps."

22 Modified-site 22 /note= "the C-terminal comprises Y-(X)-Z"

XX WO9318054-A2.
 XX 16-SEP-1993.
 XX 08-MAR-1993; 93WO-EP000517.
 XX 06-MAR-1992; 92EP-00400598.
 XX (INNO-) INNOGENETICS NV SA.
 XX De Leys R;
 XX WPI; 1993-303397/38.
 XX New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with increased antigenicity, useful in antibodies detection and vaccines against hepatitis C, HIV and HTLV.
 XX Claim 4; Page 90-98; 133pp; English.
 XX Peptide compns. comprise at least one and pref. a combination of two, three, four or more biotinylated peptides chosen from the sequences given in AAR41058-K41166. The peptides represent immunologically important regions of viral proteins and are prepd. by solid phase peptide synthesis. The compns. are useful for the detection of antibodies to HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 22 AA;

Query Match 6.8%; Score 8; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMBECSQ 67
 Db 8 DEMBECSQ 15
 |||||

RESULT 580
 AAR41107
 ID AAR41107 standard; peptide; 22 AA.
 AC AAR41107;
 XX 25-MAR-2003 (revised)
 DT 22-MAR-1994 (first entry)
 XX HCV peptide IX or HCV2 (aa 1694-1713).
 DE Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 1 /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B= biotin; X= biotinylation cpd. incorporated during synthesis; Y= bond or linking gp(s). which minimises steric hindrance, where Y is not a bond it is pref. 1-10 residues of (same or different) glycine, beta-alanine, 4-aminobutyric acid, 5-aminovaleric acid or 6-aminohexanoic acid; parentthesis around B and X indicate opt. presence

at the specified positions but B or X must be present in at least one of the positions shown, B interacts with the peptide to give a cpd. with greater diagnostic sensitivity; A (optional)= one or more amino acids, NH2 or gp. which modifies the N-terminus; Z= one or more amino acids, OH, NH2, or a linkage involving either of these 2 gps."

Modified-site 22
/note= "the C-terminal comprises Y-(X)-Z"

W09318054-A2.
16-SEP-1993.
08-MAR-1993; 93WO-EP000517.
06-MAR-1992; 92EP-00400598.
(INNO-) INNOGENETICS NV SA.
De Leys R;
WPI; 1993-303397/38.
New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with increased antigenicity, useful in antibodies detection and vaccines against hepatitis C, HIV and HTLV.
Claim 4; Page 90-98; 133pp; English.
Peptide compns. comprise at least one and pref. a combination of two, three, four or more biotinylated peptides chosen from the sequences given in AAR41058-R41166. The peptides represent immunologically important regions of viral proteins and are prepd. by solid phase peptide synthesis. The compns. are useful for the detection of antibodies to HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 22 AA;
Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||
Db 14 DEMEECSQ 21

RESULT 581
AAR41109
ID AAR41109 standard; peptide; 22 AA.
XX
AC AAR41109;
XX
25-MAR-2003 (revised)
DT 22-MAR-1994 (first entry)
XX
HCV peptide X or HCV4 (aa 1706-1725).
XX
Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
XX
OS Synthetic.
XX
Key Location/Qualifiers
FH Modified-site 1
FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B= biotin; X= biotinylation cpd. incorporated during synthesis; Y= bond or linking gp(s). which minimises steric hindrance, where Y is not a bond it is pref. 1-10 residues of (same or different) glycine, beta-alanine, 4-

aminobutyric acid, 5-aminovaleric acid or 6-aminohexanoic acid; parenthesis around B and X indicate opt. presence at the specified positions but B or X must be present in at least one of the positions shown, B interacts with the peptide to give a cpd. with greater diagnostic sensitivity; A (optional)= one or more amino acids, NH2 or gp. which modifies the N-terminus; Z= one or more amino acids, OH, NH2, or a linkage involving either of these 2 gps."

Modified-site 22
/note= "the C-terminal comprises Y-(X)-Z"

W09318054-A2.
16-SEP-1993.
08-MAR-1993; 93WO-EP000517.
06-MAR-1992; 92EP-00400598.
(INNO-) INNOGENETICS NV SA.
De Leys R;
WPI; 1993-303397/38.
New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with increased antigenicity, useful in antibodies detection and vaccines against hepatitis C, HIV and HTLV.
Claim 4; Page 90-98; 133pp; English.
Peptide compns. comprise at least one and pref. a combination of two, three, four or more biotinylated peptides chosen from the sequences given in AAR41058-R41166. The peptides represent immunologically important regions of viral proteins and are prepd. by solid phase peptide synthesis. The compns. are useful for the detection of antibodies to HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct PN field.)

Sequence 22 AA;
Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||
Db 2 DEMEECSQ 9

RESULT 582
AAR13249
ID AAR13249 standard; peptide; 22 AA.
XX
AC AAR13249;
XX
14-NOV-1997 (first entry)
DT
XX
Hepatitis C virus serotype 3 non-structural protein 4 antigen.
DE
HCV; serotype 3; non-structural protein 4; NS4; antigen; epitope;
KW immunosorbent assay; ELISA;
KW detection; discrimination.
XX
OS Hepatitis C virus.
XX
W09708198-A1.
PN
06-MAR-1997.
PD
29-AUG-1996; 96WO-JP002416.
PP
XX

```

PR 31-AUG-1995; 95JP-00223628.
XX (SRLS-) SRL INC.
XX Kumazawa T, Kiya Y, Tagami H;
XX WPI; 1997-179178/16.
XX
XX Peptide(s) antigenic for hepatitis C virus antibodies - useful for HCV
XX diagnosis, and serotype assigning assays.
XX
XX Claim 1; Page 36; 49pp; Japanese.
XX
XX The present peptide is a hepatitis C virus (HCV) serotype 3 non-
XX structural protein 4 (NS4) antigen. It can be used in immunoassays for
XX HCV antibodies, preferably an enzyme linked immunosorbant assay (ELISA).
XX Using the peptide together with other peptides antigenic for HCV
XX antibodies, e.g. corresponding to part of the NS4, core or NS5 region,
XX enables different HCV serotypes to be discriminated
XX
XX Sequence 22 AA;
    Query Match      6.8%; Score 8; DB 2; Length 22;
    Best Local Similarity 100.0%; Pred. No. 5.3;
    Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 QVIAHQPK 83
Db 11 QVIAHQPK 18

RESULT 593
AAR70220
ID AAR70220 standard; peptide; 23 AA.
XX
XX AAR70220;
XX
XX 25-MAR-2003 (revised)
XX 07-NOV-1995 (first entry)
XX
XX Hepatitis C virus (HC-J1/CDC/CHI) derived synthetic antigen.
XX
XX Composite hepatitis C virus; HC-J1/CDC/CHI; HCV; non-A non-B;
XX synthetic antigens; blood screening.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 1 /note= "H, one or more N-terminal amino acids or chemical
XX groups (used to link the peptide to a solid phase or
XX carrier)"
XX Misc-difference 22 /note= "A bond, one or more amino acids or chemical
XX groups (used to link the peptide to a solid phase or
XX carrier)"
XX Misc-difference 23 /note= "NH2, OH or a linkage involving either of these
XX groups"
XX
XX EP644202-A1.
XX
XX 22-MAR-1995.
XX
XX 14-DEC-1990; 94EP-00108611.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX WPI, 1995-116946/16.

PR 31-AUG-1995; 95JP-00223628.
XX (SRLS-) SRL INC.
XX Kumazawa T, Kiya Y, Tagami H;
XX WPI; 1997-179178/16.
XX
XX Peptide(s) antigenic for hepatitis C virus antibodies - useful for HCV
XX diagnosis, and serotype assigning assays.
XX
XX Claim 1; Page 36; 49pp; Japanese.
XX
XX The present peptide is a hepatitis C virus (HCV) serotype 3 non-
XX structural protein 4 (NS4) antigen. It can be used in immunoassays for
XX HCV antibodies, preferably an enzyme linked immunosorbant assay (ELISA).
XX Using the peptide together with other peptides antigenic for HCV
XX antibodies, e.g. corresponding to part of the NS4, core or NS5 region,
XX enables different HCV serotypes to be discriminated
XX
XX Sequence 22 AA;
    Query Match      6.8%; Score 8; DB 2; Length 22;
    Best Local Similarity 100.0%; Pred. No. 5.3;
    Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 76 QVIAHQPK 83
Db 11 QVIAHQPK 18

RESULT 593
AAR70220
ID AAR70220 standard; peptide; 23 AA.
XX
XX AAR70220;
XX
XX 25-MAR-2003 (revised)
XX 07-NOV-1995 (first entry)
XX
XX Hepatitis C virus (HC-J1/CDC/CHI) derived synthetic antigen.
XX
XX Composite hepatitis C virus; HC-J1/CDC/CHI; HCV; non-A non-B;
XX synthetic antigens; blood screening.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 1 /note= "H, one or more N-terminal amino acids or chemical
XX groups (used to link the peptide to a solid phase or
XX carrier)"
XX Misc-difference 22 /note= "A bond, one or more amino acids or chemical
XX groups (used to link the peptide to a solid phase or
XX carrier)"
XX Misc-difference 23 /note= "NH2, OH or a linkage involving either of these
XX groups"
XX
XX EP644202-A1.
XX
XX 22-MAR-1995.
XX
XX 14-DEC-1990; 94EP-00108611.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX WPI, 1995-116946/16.

XX
XX Synthetic antigens for the detection of hepatitis C virus antibodies -
XX comprise portions of the HCV peptide sequence, for use in screening blood
XX and blood products.
XX
XX Claim 4; Page 39; 51pp; English.
XX
XX AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from
XX which the synthetic HCV antigens described in AAR70210-R70229 were
XX derived. These synthetic antigens can be used to screen blood, or blood
XX products for the presence HCV, they can also be used in various specific
XX assays for the detection of HCV antibodies, and antigens, or as
XX immunogens. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-
XX MAR-2003 to correct PF field.)
XX
XX Sequence 23 AA;
    Query Match      6.8%; Score 8; DB 2; Length 23;
    Best Local Similarity 100.0%; Pred. No. 5.5;
    Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 2 DEMEECSQ 9

RESULT 584
AAR70219
ID AAR70219 standard; peptide; 23 AA.
XX
XX AAR70219;
XX
XX 25-MAR-2003 (revised)
XX 07-NOV-1995 (first entry)
XX
XX Hepatitis C virus (HC-J1/CDC/CHI) derived synthetic antigen.
XX
XX Composite hepatitis C virus; HC-J1/CDC/CHI; HCV; non-A non-B;
XX synthetic antigens; blood screening.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
XX Misc-difference 1 /note= "H, one or more N-terminal amino acids or chemical
XX groups (used to link the peptide to a solid phase or
XX carrier)"
XX Misc-difference 22 /note= "A bond, one or more amino acids or chemical
XX groups (used to link the peptide to a solid phase or
XX carrier)"
XX Misc-difference 23 /note= "NH2, OH or a linkage involving either of these
XX groups"
XX
XX EP644202-A1.
XX
XX 22-MAR-1995.
XX
XX 14-DEC-1990; 94EP-00108611.
XX
XX 14-DEC-1990; 90EP-00124241.
XX
XX (INNO-) INNOGENETICS NV.
XX
XX Deleys RJ, Pollet D, Maertens G, Van Heuverswyn H;
XX WPI, 1995-116946/16.
XX
XX Synthetic antigens for the detection of hepatitis C virus antibodies -
XX comprise portions of the HCV peptide sequence, for use in screening blood
XX and blood products.

```


PS Claim 4; Page 39; 51pp; English.

XX AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from

CC which the synthetic HCV antigens described in AAR70210-R70229 were

CC derived. These synthetic antigens can be used to screen blood, or blood

CC products for the presence of HCV, they can also be used in various specific

CC assays for the detection of HCV antibodies, and antigens, or as

CC immunogens. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-

CC MAR-2003 to correct PF field.)

XX Sequence 23 AA;

Query Match 6.8%; Score 8; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67

DB 14 DEMECSQ 21

RESULT 585

AAR34041

ID AAR34041 standard; peptide; 24 AA.

XX AAR34041;

AC AAR34041;

CC 25-MAR-2003 (revised)

DT 22-JUL-1993 (first entry)

DE HCV NS peptide DPI.

XX Hepatitis C virus; non structural region; antigen; diagnosis; vaccine.

KW Synthetic.

OS WO9306488-A1.

PN 01-APR-1993.

PD 16-SEP-1992; 92WO-US007865.

PF 16-SEP-1991; 91US-00762135.

PR 12-FEB-1992; 92US-00835717.

XX (GENE-) GENELABS TECHNOLOGIES INC.

PA Dreesman GR, Burk KH, Pauletti D;

PI WPI; 1993-117737/14.

DR Detection of hepatitis C virus antigens - using HCV reactive antibody

PT bound to solid support and competitive HCV antigen-reporter complex.

XX Example 7; Page 120; 190pp; English.

PS The synthetic peptide corresponds to residues 1694-1717 of the NS4 domain

CC of HCV. The peptide may be used for the sensitive and specific detection

CC of HCV infection. Antibodies raised against the peptide can be used for

CC passive immuno-precipitation and the HCV antigens can be used in vaccines

CC to prevent HCV infection. See also AAR34042-65. (Updated on 25-MAR-2003

CC to correct PN field.)

XX Sequence 24 AA;

Query Match 6.8%; Score 8; DB 2; Length 24;

Best Local Similarity 100.0%; Pred. No. 5.7;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67

DB 13 DEMECSQ 20

RESULT 586

AAE21849

ID AAE21849 standard; peptide; 25 AA.

XX AAE21849;

AC AAE21849;

DT 16-JUL-2002 (first entry)

DE Hepatitis C virus NS3/4A peptide #1.

XX Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;

KW virucide.

OS Hepatitis C virus.

PN WO200214362-A2.

XX 21-FEB-2002.

XX 15-AUG-2001; 2001WO-IB001774.

XX 17-AUG-2000; 2000US-0225767P.

PR 29-AUG-2000; 2000US-0229175P.

PR 03-NOV-2000; 2000US-00705547.

XX (TRIP-) TRIPEP AB.

PA Sallberg M;

PI WPI; 2002-339446/37.

DR Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or

PT absence of hepatitis C virus in a subject and for preparing a medicament

PT for treating hepatitis C virus infection.

XX Claim 20; Page 9; 90pp; English.

XX The present invention relates to novel hepatitis C virus (HCV) NS3/4A

CC proteins and their corresponding polynucleotides. NS3/4A sequences are

CC useful for identifying the presence or absence of HCV in a subject. They are

CC useful for preparing a medicament used for treating or preventing HCV

CC infection. Sequences of the invention are also used as vaccines. The

CC present sequence is HCV NS3/4A peptide

XX Sequence 25 AA;

Query Match 6.8%; Score 8; DB 5; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.9;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9

DB 6 CMSADLEV 13

RESULT 587

AAE21853

ID AAE21853 standard; peptide; 25 AA.

XX AAE21853;

AC AAE21853;

DT 16-JUL-2002 (first entry)

DE Hepatitis C virus NS3/4A mutant peptide #3.

XX Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;

KW virucide; mutant; mutein.

XX Hepatitis C virus.

OS Synthetic.

XX WO200214362-A2.

```

XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M;
XX DR WPI; 2002-339446/37.
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX PT absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Claim 20; Page 9; 90pp; English.
XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX CC proteins and their corresponding polynucleotides. NS3/4A sequences are
XX CC useful for identifying the presence or absence HCV in a subject. They are
XX CC useful for preparing a medicament used for treating or preventing HCV
XX CC infection. Sequences of the invention are also used as vaccines. The
XX CC present sequence is HCV NS3/4A mutant peptide in which the proteolytic
XX CC cleavage site is altered
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 589
AAE21855
ID AAE21855 standard; peptide; 25 AA.
XX AC AAE21855;
XX DT 16-JUL-2002 (first entry)
XX DE Hepatitis C virus NS3/4A mutant peptide #6.
XX KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX KW virucide; mutant; mutein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN WO200214362-A2.
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M;
XX DR WPI; 2002-339446/37.
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX PT absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Claim 20; Page 9; 90pp; English.
XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX CC proteins and their corresponding polynucleotides. NS3/4A sequences are
XX CC useful for identifying the presence or absence HCV in a subject. They are
XX CC useful for preparing a medicament used for treating or preventing HCV
XX CC infection. Sequences of the invention are also used as vaccines. The
XX CC present sequence is HCV NS3/4A mutant peptide in which the proteolytic
XX CC cleavage site is altered
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 588
AAE21855
ID AAE21855 standard; peptide; 25 AA.
XX AC AAE21855;
XX DT 16-JUL-2002 (first entry)
XX DE Hepatitis C virus NS3/4A mutant peptide #5.
XX KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX KW virucide; mutant; mutein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN WO200214362-A2.
XX PD 21-FEB-2002.
XX PF 15-AUG-2001; 2001WO-IB001774.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX PA (TRIP-) TRIPEP AB.
XX PI Sallberg M;
XX DR WPI; 2002-339446/37.
XX PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or

```

Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 590
AAE21857
ID AAE21857 standard; peptide; 25 AA.
AC AAE21857;
XX
DT 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A mutant peptide #7.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutain.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX
PI Sallberg M;
XX
DR WPI; 2002-339446/37.
XX
PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
PS Claim 20; Page 9; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant peptide in which the proteolytic
CC cleavage site is altered
XX
SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 592
AAE21852
ID AAE21852 standard; peptide; 25 AA.
XX
AC AAE21852;
XX
DT 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A mutant peptide #2.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutain.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.

Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 591
AAE21858
ID AAE21858 standard; peptide; 25 AA.
AC AAE21858;
XX
DT 16-JUL-2002 (first entry)
XX

PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M;
XX WPI; 2002-339446/37.
DR
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
XX Claim 20; Page 9; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant peptide in which the proteolytic
CC cleavage site is altered
XX
SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||
AE21851
ID AAE21851 standard; peptide; 25 AA.
XX
XX AC AAE21851;
XX
XX 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A mutant peptide #1.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX WO200214362-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-IB001774.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 03-NOV-2000; 2000US-00705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M;
XX
XX WPI; 2002-339446/37.
XX
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
XX Claim 20; Page 9; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are
CC useful for identifying the presence or absence of HCV in a subject. They are
CC useful for preparing a medicament used for treating or preventing HCV
CC infection. Sequences of the invention are also used as vaccines. The
CC present sequence is HCV NS3/4A mutant peptide in which the proteolytic
CC cleavage site is altered
XX
SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||
AE21854
ID AAE21854 standard; peptide; 25 AA.
XX
XX AC AAE21854;
XX
XX 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A mutant peptide #4.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; mutant; mutein.
XX
OS Hepatitis C virus.
OS Synthetic.
XX
XX WO200214362-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-IB001774.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 03-NOV-2000; 2000US-00705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Sallberg M;
XX
XX WPI; 2002-339446/37.
XX
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
PT absence of hepatitis C virus in a subject and for preparing a medicament
PT for treating hepatitis C virus infection.
XX
XX Claim 20; Page 9; 90pp; English.
XX
CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
CC proteins and their corresponding polynucleotides. NS3/4A sequences are

```
RESULT 595
AAE19903
ID AAE19903 standard; peptide; 25 AA.
XX
XX AC AAE19903;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Hepatitis C virus (HCV) NS3/4A mutant peptide #2.
XX
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX cytotostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
XX mutant; mutein.
XX
XX OS Hepatitis C virus.
XX Synthetic.
XX
XX PN WO200213855-A2.
XX
XX PD 21-FEB-2002.
XX
XX PF 15-AUG-2001; 2001WO-IB001808.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX
XX PR 29-AUG-2000; 2000US-0229175P.
XX
XX PR 03-NOV-2000; 2000US-00705547.
XX
XX PA (TRIP-) TRIPEP AB.
XX
XX PI Sallberg M, Hultgren C;
XX
XX DR WPI; 2002-241837/29.
XX
XX PS Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Example 6; Page 25; 120pp; English.
XX
XX CC The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to a
XX hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A mutant peptide
XX
XX SQ Sequence 25 AA;
XX
XX Query Match 6.8%; Score 8; DB 5; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 CMSADLEV 9
XX |||||
XX Db 6 CMSADLEV 13
XX
XX RESULT 596
XX AAE19912
XX ID AAE19912 standard; peptide; 25 AA.
XX
XX AC AAE19912;
XX
XX XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Hepatitis C virus (HCV) NS3/4A mutant peptide #6.
XX
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX cytotostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
XX mutant; mutein.
XX
```

```
XX
XX OS Hepatitis C virus.
XX Synthetic.
XX
XX PN WO200213855-A2.
XX
XX XX
XX PD 21-FEB-2002.
XX
XX XX
XX PF 15-AUG-2001; 2001WO-IB001808.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX
XX PR 29-AUG-2000; 2000US-0229175P.
XX
XX PR 03-NOV-2000; 2000US-00705547.
XX
XX PA (TRIP-) TRIPEP AB.
XX
XX PI Sallberg M, Hultgren C;
XX
XX XX
XX DR WPI; 2002-241837/29.
XX
XX XX
XX PT Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX PS Example 6; Page 25; 120pp; English.
XX
XX CC The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to a
XX hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is HCV NS3/4A mutant peptide
XX
XX SQ Sequence 25 AA;
XX
XX Query Match 6.8%; Score 8; DB 5; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 2 CMSADLEV 9
XX |||||
XX Db 6 CMSADLEV 13
XX
XX RESULT 597
XX AAE19913
XX ID AAE19913 standard; peptide; 25 AA.
XX
XX AC AAE19913;
XX
XX XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Hepatitis C virus (HCV) NS3/4A mutant peptide #7.
XX
XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX cytotostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
XX mutant; mutein.
XX
XX OS Hepatitis C virus.
XX Synthetic.
XX
XX PN WO200213855-A2.
XX
XX XX
XX PD 21-FEB-2002.
XX
XX PF 15-AUG-2001; 2001WO-IB001808.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX
XX PR 29-AUG-2000; 2000US-0229175P.
XX
XX PR 03-NOV-2000; 2000US-00705547.
XX
XX XX
```

PA (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 25; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A mutant peptide
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
RESULT 598
ID AAE19914 standard; peptide; 25 AA.
XX AAE19914;
XX 18-JUN-2002 (first entry)
XX Hepatitis C virus (HCV) NS3/4A mutant peptide #8.
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW mutant; mutain.
XX Hepatitis C virus.
OS Synthetic.
XX WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001808.
XX 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 25; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A mutant peptide
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
RESULT 598
ID AAE19914 standard; peptide; 25 AA.
XX AAE19914;
XX 18-JUN-2002 (first entry)
XX Hepatitis C virus (HCV) NS3/4A mutant peptide #8.
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW mutant; mutain.
XX Hepatitis C virus.
OS Synthetic.
XX WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001808.
XX 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 25; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A mutant peptide
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
RESULT 599
ID AAE19901 standard; peptide; 25 AA.
XX AAE19901;
XX 18-JUN-2002 (first entry)
XX Hepatitis C virus (HCV) NS3/4A peptide #1.
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX Hepatitis C virus.
OS WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001808.
XX 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 25; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A peptide
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A mutant peptide
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
RESULT 599
ID AAE19901 standard; peptide; 25 AA.
XX AAE19901;
XX 18-JUN-2002 (first entry)
XX Hepatitis C virus (HCV) NS3/4A peptide #1.
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX Hepatitis C virus.
OS WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001808.
XX 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Example 6; Page 25; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A peptide
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

```
RESULT 600
AAE19909
ID AAE19909 standard; peptide; 25 AA.
AC AAE19909;
XX
XX
DT 18-JUN-2002 (first entry)
XX
XX Hepatitis C virus (HCV) NS3/4A mutant peptide #3.
DE
XX
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW mutant; mutein.
XX
XX Hepatitis C virus.
OS Synthetic.
OS
XX WO200213855-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 15-AUG-2001; 2001WO-IB001808.
PF
XX
XX 17-AUG-2000; 2000US-0225767P.
PR
XX 29-AUG-2000; 2000US-0229175P.
PR
XX 03-NOV-2000; 2000US-00705547.
PR
XX (TRIP-) TRIPEP AB.
PA
XX
XX Sallberg M, Hultgren C;
PI
XX WPI; 2002-241837/29.
DR
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Example 6; Page 25; 120pp; English.
PS
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3/4A mutant peptide
XX
XX Sequence 25 AA;
SQ
Query Match 6.8%; Score 8; DB 5; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
RESULT 602
AAE19902
ID AAE19902 standard; peptide; 25 AA.
XX
XX AAE19902;
AC
XX
XX 18-JUN-2002 (first entry)
DT
XX
XX Hepatitis C virus (HCV) NS3/4A mutant peptide #1.
DE
XX
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW mutant; mutein.
XX
XX Hepatitis C virus.
OS Synthetic.
OS
XX WO200213855-A2.
PN
XX
XX 21-FEB-2002.
PD
XX
XX 15-AUG-2001; 2001WO-IB001808.
PF
XX
XX 17-AUG-2000; 2000US-0225767P.
PR
XX 29-AUG-2000; 2000US-0229175P.
PR
XX 03-NOV-2000; 2000US-00705547.
PR
```

```
RESULT 601
AAE19911
ID AAE19911 standard; peptide; 25 AA.
XX
XX AAE19911;
AC
XX
XX 18-JUN-2002 (first entry)
DT
XX
XX Hepatitis C virus (HCV) NS3/4A mutant peptide #5.
DE
XX
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
KW cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer;
KW
```

XX PA (TRIP-) TRIPEP AB.
 XX PI Sallberg M, Hultgren C;
 XX DR WPI; 2002-241837/29.
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Example 6; Page 25; 120pp; English.
 XX The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is HCV NS3/4A mutant peptide
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 6.8%; Score 8; DB 5; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CMSADLEV 9
 DB |||||
 6 CMSADLEV 13
 RESULT 603
 AAE19910
 ID AAE19910 standard; peptide; 25 AA.
 XX AC AAE19910;
 XX 18-JUN-2002 (first entry)
 XX Hepatitis C virus (HCV) NS3/4A mutant peptide #4.
 XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
 KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
 KW mutant; mutin.
 XX Hepatitis C virus.
 OS Synthetic.
 XX WO200213855-A2.
 XX 21-FEB-2002.
 XX 15-AUG-2001; 2001WO-IB001808.
 XX 17-AUG-2000; 2000US-0225767P.
 XX 29-AUG-2000; 2000US-0229175P.
 XX 03-NOV-2000; 2000US-00705547.
 XX (TRIP-) TRIPEP AB.
 XX Sallberg M, Hultgren C;
 XX WPI; 2002-241837/29.
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Example 6; Page 25; 120pp; English.
 XX The invention relates to a composition comprising ribavirin and an

CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is HCV NS3/4A mutant peptide
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 6.8%; Score 8; DB 5; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CMSADLEV 9
 DB |||||
 6 CMSADLEV 13
 RESULT 604
 ABW00362
 ID ABW00362 standard; peptide; 25 AA.
 XX AC ABW00362;
 XX 15-JAN-2004 (first entry)
 XX Hepatitis C virus NS3/4A mutant peptide #5.
 XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
 KW virucide; mutant; mutin.
 XX Hepatitis C virus.
 OS Synthetic.
 XX US2002136740-A1.
 XX 26-SEP-2002.
 XX 15-AUG-2001; 2001US-00929955.
 XX 17-AUG-2000; 2000US-0225767P.
 XX 29-AUG-2000; 2000US-0229175P.
 XX (SALL/) SALLBERG M.
 XX (HULT/) HULTGREN C.
 XX Sallberg M, Hultgren C;
 XX WPI; 2003-764978/72.
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Example 6; Page 13; 0pp; English.
 XX The invention relates to a composition comprising ribavirin and an
 CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
 CC is useful in enhancing the immune response to a hepatitis C antigen where
 CC the composition is delivered to an animal identified as requiring an
 CC enhanced immune response. The vaccine is useful in the treatment and
 CC prevention of hepatitis C infection. The present sequence is Hepatitis C
 CC virus NS3/4A mutant peptide
 XX
 XX Sequence 25 AA;
 SQ
 Query Match 6.8%; Score 8; DB 7; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 CMSADLEV 9
 DB |||||


```
XX DR WPI; 2003-764978/72.
XX PT Vaccine compositions for treating and preventing disease, preferably
XX PT hepatitis C virus infection, comprises ribavirin and antigen that has
XX PT epitope present in hepatitis C virus.
XX PS Example 6; Page 13; Opp; English.
XX CC The invention relates to a composition comprising ribavirin and an
XX CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
XX CC is useful in enhancing the immune response to a hepatitis C antigen where
XX CC the composition is delivered to an animal identified as requiring an
XX CC enhanced immune response. The vaccine is useful in the treatment and
XX CC prevention of hepatitis C infection. The present sequence is Hepatitis C
XX CC virus NS3/4A mutant peptide
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||

RESULT 608
ABW00364
ID ABW00364 standard; peptide; 25 AA.
XX AC ABW00364;
XX DT 15-JAN-2004 (first entry)
XX DE Hepatitis C virus NS3/4A mutant peptide #2.
XX KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX KW virucide; mutant; mutein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN US2002136740-A1.
XX PD 26-SEP-2002.
XX PF 15-AUG-2001; 2001US-00929955.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PA (SALL/) SALLBERG M.
XX PA (HULT/) HULTGREN C.
XX PI Sallberg M, Hultgren C;
XX PI WPI; 2003-764978/72.
XX DR Vaccine compositions for treating and preventing disease, preferably
XX DR hepatitis C virus infection, comprises ribavirin and antigen that has
XX DR epitope present in hepatitis C virus.
XX PS Example 6; Page 13; Opp; English.
XX CC The invention relates to a composition comprising ribavirin and an
XX CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
XX CC is useful in enhancing the immune response to a hepatitis C antigen where
XX CC the composition is delivered to an animal identified as requiring an
XX CC enhanced immune response. The vaccine is useful in the treatment and
XX CC prevention of hepatitis C infection. The present sequence is Hepatitis C
XX CC virus NS3/4A mutant peptide
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||

RESULT 608
ABW00364
ID ABW00364 standard; peptide; 25 AA.
XX AC ABW00364;
XX DT 15-JAN-2004 (first entry)
XX DE Hepatitis C virus NS3/4A mutant peptide #7.
XX KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX KW virucide; mutant; mutein.
XX OS Hepatitis C virus.
XX OS Synthetic.
XX PN US2002136740-A1.
XX PD 26-SEP-2002.
XX PF 15-AUG-2001; 2001US-00929955.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PA (SALL/) SALLBERG M.
XX PA (HULT/) HULTGREN C.
XX PI Sallberg M, Hultgren C;
XX PI WPI; 2003-764978/72.
XX DR Vaccine compositions for treating and preventing disease, preferably
XX DR hepatitis C virus infection, comprises ribavirin and antigen that has
XX DR epitope present in hepatitis C virus.
XX PS Example 6; Page 13; Opp; English.
XX CC The invention relates to a composition comprising ribavirin and an
XX CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
XX CC is useful in enhancing the immune response to a hepatitis C antigen where
XX CC the composition is delivered to an animal identified as requiring an
XX CC enhanced immune response. The vaccine is useful in the treatment and
XX CC prevention of hepatitis C infection. The present sequence is Hepatitis C
XX CC virus NS3/4A mutant peptide
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||

RESULT 610
ABW00365
ID ABW00365 standard; peptide; 25 AA.
XX
```


XX CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A mutant peptide
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||

RESULT 613
ADG47679
ID ADG47679 standard; peptide; 25 AA.
XX AC ADG47679;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #7.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX PT New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.
XX PS Example 1; SEQ ID NO 22; 83pp; English.
XX CC The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
XX The nucleic acid and the peptide enhance an immune response to a
XX hepatitis C antigen and are potent immunogens. The present sequence is
XX used in the exemplification of the invention.
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||

RESULT 615
ADG47673
ID ADG47673 standard; peptide; 25 AA.
XX AC ADG47673;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #1.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX OS Hepatitis C virus.

Db 6 CMSADLEV 13

RESULT 614
ADG47678
ID ADG47678 standard; peptide; 25 AA.
XX AC ADG47678;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #6.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX PT New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.
XX PS Example 1; SEQ ID NO 21; 83pp; English.
XX CC The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
XX The nucleic acid and the peptide enhance an immune response to a
XX hepatitis C antigen and are potent immunogens. The present sequence is
XX used in the exemplification of the invention.
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13
|||||

RESULT 615
ADG47673
ID ADG47673 standard; peptide; 25 AA.
XX AC ADG47673;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #1.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX OS Hepatitis C virus.

XX US2003206919-A1.
XX 06-NOV-2003.
XX 26-NOV-2002; 2002US-00307047.
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX (SALL/) SALLBERG M.
XX Sallberg M;
XX WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.
XX Example 1; SEQ ID NO 16; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
XX The nucleic acid and the peptide enhance an immune response to a
XX hepatitis C antigen and are potent immunogens. The present sequence is
XX used in the exemplification of the invention.
XX Sequence 25 AA;
XX Query Match 6.8%; Score 8; DB 8; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 CMSADLEV 9
XX |||||
XX 6 CMSADLEV 13
XX RESULT 616
XX ADG47675
XX ID ADG47675 standard; peptide; 25 AA.
XX AC ADG47675;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #3.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX Query Match 6.8%; Score 8; DB 8; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 CMSADLEV 9
XX |||||
XX 6 CMSADLEV 13
XX RESULT 616
XX ADG47675
XX ID ADG47675 standard; peptide; 25 AA.
XX AC ADG47675;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #3.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX Query Match 6.8%; Score 8; DB 8; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 CMSADLEV 9
XX |||||
XX 6 CMSADLEV 13
XX RESULT 617
XX ADG47676
XX ID ADG47676 standard; peptide; 25 AA.
XX AC ADG47676;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #4.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.
XX Example 1; SEQ ID NO 18; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
XX The nucleic acid and the peptide enhance an immune response to a
XX hepatitis C antigen and are potent immunogens. The present sequence is
XX used in the exemplification of the invention.
XX Sequence 25 AA;
XX Query Match 6.8%; Score 8; DB 8; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 CMSADLEV 9
XX |||||
XX 6 CMSADLEV 13
XX RESULT 617
XX ADG47676
XX ID ADG47676 standard; peptide; 25 AA.
XX AC ADG47676;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #4.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.
XX Example 1; SEQ ID NO 19; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
XX hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
XX The nucleic acid and the peptide enhance an immune response to a
XX hepatitis C antigen and are potent immunogens. The present sequence is
XX used in the exemplification of the invention.
XX Sequence 25 AA;
XX Query Match 6.8%; Score 8; DB 8; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 5.9;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 CMSADLEV 9
XX |||||
XX 6 CMSADLEV 13

Query Match

6.8%; Score 8; DB 8; Length 25;

```
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 618
ADG47674
ID ADG47674 standard; peptide; 25 AA.
XX AC ADG47674;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #2.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
XX PI Sallberg M;
XX DR WPI; 2004-051480/05.
XX PT New purified or isolated nucleic acid useful for enhancing an immune
XX PT response to a hepatitis C antigen comprises specific nucleotide sequences
XX PT and the amino acid sequences.
XX PS Example 1; SEQ ID NO 17; 83pp; English.
XX PS The invention relates to a purified or isolated nucleic acid. The
XX CC peptides are useful as immunogens for the treatment and prevention of
XX CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
XX CC The nucleic acid and the peptide enhance an immune response to a
XX CC hepatitis C antigen and are potent immunogens. The present sequence is
XX CC used in the exemplification of the invention.
XX SQ Sequence 25 AA;

Query Match 6.8%; Score 8; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 619
ADG47671
ID ADG47671 standard; peptide; 25 AA.
XX AC ADG47671;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain peptide #1.

Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 620
ADG47677
ID ADG47677 standard; peptide; 25 AA.
XX AC ADG47677;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #5.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.

immunogen; hepatitis C virus; HCV infection; vaccine.
Hepatitis C virus.
US2003206919-A1.
06-NOV-2003.
26-NOV-2002; 2002US-00307047.
17-AUG-2000; 2000US-0225767P.
29-AUG-2000; 2000US-0229175P.
15-AUG-2001; 2001US-00929955.
15-AUG-2001; 2001US-00930591.
(SALL/) SALLBERG M.
Sallberg M;
WPI; 2004-051480/05.
New purified or isolated nucleic acid useful for enhancing an immune
response to a hepatitis C antigen comprises specific nucleotide sequences
and the amino acid sequences.
Example 1; SEQ ID NO 14; 83pp; English.
The invention relates to a purified or isolated nucleic acid. The
peptides are useful as immunogens for the treatment and prevention of
hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
The nucleic acid and the peptide enhance an immune response to a
hepatitis C antigen and are potent immunogens. The present sequence is
used in the exemplification of the invention.
Sequence 25 AA;

Query Match 6.8%; Score 8; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 620
ADG47677
ID ADG47677 standard; peptide; 25 AA.
XX AC ADG47677;
XX DT 11-MAR-2004 (first entry)
XX DE HCV NS3/4A domain mutant peptide #5.
XX KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutain.
XX OS Synthetic.
XX OS Hepatitis C virus.
XX PN US2003206919-A1.
XX PD 06-NOV-2003.
XX PF 26-NOV-2002; 2002US-00307047.
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 15-AUG-2001; 2001US-00929955.
XX PR 15-AUG-2001; 2001US-00930591.
XX PA (SALL/) SALLBERG M.
```

XX Sallberg M;
PI WPI; 2004-051480/05.
XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX Example 1; SEQ ID NO 20; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.
XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CMSADLEV 9
DB 6 CMSADLEV 13
|||||
RESULT 621
ADG47680
ID ADG47680 standard; peptide; 25 AA.
XX
AC ADG47680;
XX
DT 11-MAR-2004 (first entry)
XX
DE HCV NS3/4A domain mutant peptide #8.
XX immunogen; hepatitis C virus; HCV infection; vaccine; mutant; muten.
XX Synthetic.
OS Hepatitis C virus.
XX
XX US2003206919-A1.
XX
XX 06-NOV-2003.
XX
XX 26-NOV-2002; 2002US-00307047.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 15-AUG-2001; 2001US-00929955.
XX 15-AUG-2001; 2001US-00930591.
XX
XX (SALL/) SALLBERG M.
XX
XX Sallberg M;
XX
XX WPI; 2004-051480/05.
XX
XX New purified or isolated nucleic acid useful for enhancing an immune
XX response to a hepatitis C antigen comprises specific nucleotide sequences
XX and the amino acid sequences.
XX Example 1; SEQ ID NO 23; 83pp; English.
XX The invention relates to a purified or isolated nucleic acid. The
XX peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
CC used in the exemplification of the invention.

XX
SQ Sequence 25 AA;
Query Match 6.8%; Score 8; DB 8; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CMSADLEV 9
DB 6 CMSADLEV 13
|||||
RESULT 622
AAR41183
ID AAR41183 standard; peptide; 26 AA.
XX
AC AAR41183;
XX
XX 25-MAR-2003 (revised)
DT 22-MAR-1994 (first entry)
XX
XX HCV NS4 protein HCV4/5.
XX Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Region 12..19
FT /label= epitope_2B
XX
XX WO9318054-A2.
XX
XX PD 16-SEP-1993.
XX
XX PF 08-MAR-1993; 93WO-EP000517.
XX
XX PR 06-MAR-1992; 92EP-00400598.
XX
XX PA (INNO-) INNOGENETICS NV SA.
XX
XX PI De Leys R;
XX
XX WPI; 1993-303397/38.
XX
XX New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
XX increased antigenicity, useful in antibodies detection and vaccines
XX against hepatitis C, HIV and HTLV.
XX Disclosure; Page 79; 133pp; English.
XX
XX Peptide compsns. comprise at least one and pref. a combination of two,
XX three, four or more biotinylated peptides chosen from the sequences given
XX in AAR41058-R41166. The peptides may be hybrids consisting of
XX combinations of the core epitopes of the HCV core (AAR41171-R41180), HCV
XX NS4 (AAR41181-R41186) or the HCV NS5 (AAR41187-R41193) region separated
XX by Gly and/or Ser residues. Pref. hybrid peptides are given in AAR41161-
XX R41163. The peptides represent immunologically important regions of viral
XX proteins and are prep'd. by solid phase peptide synthesis. The compsns.
XX are useful for the detection of antibodies to HCV, and/or HIV, and/or
XX HTLV-I or II. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 26 AA;
Query Match 6.8%; Score 8; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMECSQ 67
DB 1 DEMECSQ 8
|||||

```

RESULT 623
AAR41181
ID AAR41181 standard; peptide; 26 AA.
AC AAR41181;
XX
XX
DT 25-MAR-2003 (revised)
DT 22-MAR-1994 (first entry)
XX
XX
DE HCV NS4 protein HCV1/2.
XX
KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 9..14
FT /label= epitope_1
XX
XX WO9318054-A2.
XX
PD 16-SEP-1993.
XX
XX 08-MAR-1993; 93WO-EP000517.
XX
XX 06-MAR-1992; 92EP-00400598.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
XX De Leys R;
PI
XX
XX WPI; 1993-303397/38.
XX
XX New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
XX increased antigenicity, useful in antibodies detection and vaccines
XX against hepatitis C, HIV and HTLV.
XX
XX Disclosure; Page 79; 133pp; English.
XX
XX Peptide compans. comprise at least one and pref. a combination of two,
XX three, four or more biotinylated peptides chosen from the sequences given
XX in AAR41058-R41166. The peptides may be hybrids consisting of
XX combinations of the core epitopes of the HCV core (AAR41171-R41180), HCV
XX NS4 (AAR41181-R41186) or the HCV NS5 (AAR41187-R41193) region separated
XX by Gly and/or Ser residues. Pref. hybrid peptides are given in AAR41161-
XX R41163. The peptides represent immunologically important regions of viral
XX proteins and are prepd. by solid phase peptide synthesis. The compans.
XX are useful for the detection of antibodies to HCV, and/or HIV, and/or
XX HTLV-I or II. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 26 AA;
Query Match 6.8%; Score 8; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 19 DEMEECSQ 26
RESULT 624
AAR67297
ID AAR67297 standard; peptide; 30 AA.
XX
XX AAR67297;
XX
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
XX
XX
01-AUG-1995 (first entry)
Hepatitis C virus non-structural region 4 type dependent seq 1a.
Hepatitis C virus; HCV; genotyping; serotyping;
non-structural region 4 type dependent sequences.
Hepatitis C virus.
WO9427153-A1.
24-NOV-1994.
09-MAY-1994; 94WO-US005151.
10-MAY-1993; 93US-00060400.
(CHIR ) CHIRON CORP.
Chien DY, Kuo G;
WPI; 1995-006972/01.
Method for typing hepatitis C virus - by genotype or serotype.
Example 5; Page 26; 47pp; English.
AAR67296-R67299 are hepatitis C virus (HCV) non-structural region 4 type
dependent sequences, they react with common epitopes found between amino
acid residues 1689-1718 of HCV-1. They can be used as part of a new
method for typing HCV's by genotype or serotype. (Updated on 25-MAR-2003
to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)
Sequence 30 AA;
Query Match 6.8%; Score 8; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 18 DEMEECSQ 25
RESULT 625
AAU84711
ID AAU84711 standard; peptide; 30 AA.
XX
XX AAU84711;
AC AAU84711;
XX
XX 08-MAY-2002 (first entry)
XX
XX HCV HepCia segment 114.
XX
XX Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
XX viral infection; human immunodeficiency virus; melanoma;
XX bacterial infection; Salmonella; Legionella; parasitic infection;
XX Trypanosoma; Toxoplasma; Giardia.
XX
XX Hepatitis C virus.
OS
XX
XX WO200190197-A1.
XX
XX 29-NOV-2001.
XX
XX 25-MAY-2001; 2001WO-AU000622.
XX
XX 26-MAY-2000; 2000AU-00007761.
XX
XX (AUSU ) UNIV AUSTRALIAN NAT.
XX
XX Thomson SA, Ramshaw IA;
XX

```


DR WPI; 2002-147575/19.
 DR N-PSDB; ABK36549.
 XX
 PT New synthetic polypeptides having several different segments of at least
 PT one parent polypeptide linked together differently compared to the
 PT linkage in the parent polypeptide, for inducing immune response against a
 PT pathogen or cancer.
 XX
 PS Example 2; Fig 26; 364pp; English.
 XX
 CC The invention relates to a new synthetic polypeptide (I) comprising
 CC several different segments of at least one parent polypeptide linked
 CC together in a different relationship relative to their linkage in the
 CC parent polypeptide to impede, abrogate or otherwise alter at least one
 CC function associated with the parent polypeptide and for inducing an
 CC immune response against a pathogen or cancer. Also included are a
 CC synthetic polynucleotide encoding and a computer system for designing the
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
 CC are referred to as a Savine. The synthetic polypeptide is useful for
 CC modulating immune responses preferably directed against a pathogen or a
 CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
 CC oesophagus, brain, testicle, uterus), as potentiating agents.
 CC Compositions comprising the polypeptide may be used in the treatment or
 CC prophylaxis against viral (such as infections caused by HIV (human
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
 CC a peptide derived from a parent protein used to construct a savine of the
 CC invention
 XX
 SQ Sequence 30 AA;
 Query Match 6.8%; Score 8; DB 5; Length 30;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 60 DEMEECSQ 67
 |||||
 Db 13 DEMEECSQ 20
 |||||
 RESULT 626
 AAU84707
 ID AAU84707 standard; peptide; 30 AA.
 AC AAU84707;
 DT 08-MAY-2002 (first entry)
 DE HCV HepC1a segment 110.
 XX
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
 KW viral infection; human immunodeficiency virus; melanoma;
 KW bacterial infection; Salmonella; Legionella; parasitic infection;
 KW Trypanosoma; Toxoplasma; Giardia.
 XX
 OS Hepatitis C virus.
 XX
 FN WO200190197-A1.
 XX
 PD 29-NOV-2001.
 XX
 PP 25-MAY-2001; 2001WO-AU000622.
 XX
 PR 26-MAY-2000; 2000AU-00007761.
 XX
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Thomson SA, Ramshaw IA;

XX WPI; 2002-147575/19.
 DR N-PSDB; ABK36545.
 XX
 PT New synthetic polypeptides having several different segments of at least
 PT one parent polypeptide linked together differently compared to the
 PT linkage in the parent polypeptide, for inducing immune response against a
 PT pathogen or cancer.
 XX
 PS Example 2; Fig 26; 364pp; English.
 XX
 CC The invention relates to a new synthetic polypeptide (I) comprising
 CC several different segments of at least one parent polypeptide linked
 CC together in a different relationship relative to their linkage in the
 CC parent polypeptide to impede, abrogate or otherwise alter at least one
 CC function associated with the parent polypeptide and for inducing an
 CC immune response against a pathogen or cancer. Also included are a
 CC synthetic polynucleotide encoding and a computer system for designing the
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
 CC are referred to as a Savine. The synthetic polypeptide is useful for
 CC modulating immune responses preferably directed against a pathogen or a
 CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
 CC oesophagus, brain, testicle, uterus), as potentiating agents.
 CC Compositions comprising the polypeptide may be used in the treatment or
 CC prophylaxis against viral (such as infections caused by HIV (human
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
 CC Salmonella, Streptococcal, Legionella and Mycobacterium or parasitic
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
 CC a peptide derived from a parent protein used to construct a savine of the
 CC invention
 XX
 SQ Sequence 30 AA;
 Query Match 6.8%; Score 8; DB 5; Length 30;
 Best Local Similarity 100.0%; Pred. No. 6.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 CMSADLEV 9
 |||||
 Db 15 CMSADLEV 22
 |||||
 RESULT 627
 AAR41182
 ID AAR41182 standard; peptide; 32 AA.
 AC AAR41182;
 DT 25-MAR-2003 (revised)
 DT 22-MAR-1994 (first entry)
 XX
 DE HCV NS4 protein HCV3/4/5.
 XX
 KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
 KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
 KW epitope; antibody; biotin; diagnosis; detection; vaccine.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 15..20
 FT /label= epitope_2A
 XX
 FN WO9318054-A2.
 XX
 PD 16-SEP-1993.
 XX
 PP 08-MAR-1993; 93WO-EP000517.
 XX

```
PR 06-MAR-1992; 92EP-00400598.
XX (INNO-) INNOGENETICS NV SA.
XX De Leys R;
XX WPI; 1993-303397/38.
XX New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
PT increased antigenicity, useful in antibodies detection and vaccines
PT against hepatitis C, HIV and HTLV.
XX Disclosure; Page 79; 133pp; English.
XX Peptide compens. comprise at least one and pref. a combination of two,
CC three, four or more biotinylated peptides chosen from the sequences given
CC in AAR41058-R41166. The peptides may be hybrids consisting of
CC combinations of the core epitopes of the HCV core (AAR41171-R41180), HCV
CC NS4 (AAR41181-R41186) or the HCV NS5 (AAR41187-R41193) region separated
CC by Gly and/or Ser residues. Pref. hybrid peptides are given in AAR41161-
CC R41163. The peptides represent immunologically important regions of viral
CC proteins and are prep'd. by solid phase peptide synthesis. The compens.
CC are useful for the detection of antibodies to HCV, and/or HIV, and/or
CC HTLV-I or II. (Updated on 25-MAR-2003 to correct PN field.)
XX Sequence 32 AA;
SQ Query Match 6.8%; Score 8; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 7.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db |||||||
7 DEMEECSQ 14

RESULT 628
AAR13766
ID AAR13766 standard; protein; 39 AA.
XX
AC AAR13766;
XX
XX 25-MAR-2003 (revised)
DT 22-NOV-1991 (first entry)
XX
XX HCV antibody detecting peptide (III).
XX Hepatitis C virus; assay; non-A non-B hepatitis; NANBH; diagnosis.
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Peptide 1. .18
XX /label= claim 10
XX Peptide 9. .39
XX /label= claim 9
XX Peptide 9. .18
XX /label= claim 1(c)
XX
XX EP445801-A.
XX
XX 11-SEP-1991.
XX
XX 07-MAR-1991; 91EP-00103471.
XX
XX 08-MAR-1990; 90JP-00058700.
XX 16-MAR-1990; 90JP-00067439.
XX 27-MAR-1990; 90JP-00080100.
XX 31-OCT-1990; 90JP-00296899.
XX
XX (KURS ) KURARAY CO LTD.
XX
XX Arima T, Namba T, Tsuji M, Tadashi H;
```

```
XX WPI; 1991-268671/37.
XX Peptide(s) specific for antibody to non-A and non-B hepatitis antigen -
PT uses as assay reagents for anti-HCV antibody for diagnosis of HCV
PT infection.
XX Claim 1(c),4,9,10; Page 48; 63pp; English.
XX The peptide may be prep'd. by solid phase synthesis or liq. phase
CC synthesis or by fragment condensan. The peptide can be used as anti-HCV
CC antibody assay reagent with high sensitivity. See also AAR13764-73 and
CC AAR13565-67. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25
CC -MAR-2003 to correct PI field.)
XX Sequence 39 AA;
SQ Query Match 6.8%; Score 8; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db |||||||
22 DEMEECSQ 29

RESULT 629
AAR25134
ID AAR25134 standard; protein; 40 AA.
XX
AC AAR25134;
XX
XX 23-DEC-1992 (first entry)
DT
XX moka-C33.
DE
XX Hepatitis C virus; blood transfusion.
KW
XX Synthetic.
OS
XX JP04159298-A.
PN
XX 02-JUN-1992.
PD
XX 19-OCT-1990; 90JP-00282431.
PF
XX 19-OCT-1990; 90JP-00282431.
PR
XX (OLYU ) OLYMPUS OPTICAL CO LTD.
XX
XX WPI; 1992-231947/28.
XX
XX New peptides acting as antigenic analogues of human hepatitis C virus -
PT useful for detecting HCV antibody positive patients.
XX
XX Claim 1; Page 1; 14pp; Japanese.
PS
XX The sequences given in AAR25130-35 are peptides which have reactivity to
CC the antibody against hepatitis C virus (HCV). They can be used on their
CC own or as a mixture two different peptides. Using these peptides, HCV
CC antibody positive patients can be detected and hepatitis caused by blood
CC transfusion can be prevented
XX Sequence 40 AA;
SQ Query Match 6.8%; Score 8; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db |||||||
21 DEMEECSQ 28
```

```

RESULT 630
AAR14078
ID AAR14078 standard; peptide; 42 AA.
XX
AC AAR14078;
XX
DT 03-DEC-1991 (first entry)
XX
DE Peptide comprising Hepatitis C virus epitope.
XX
KW anti-HCV antibody; non-A, non-B hepatitis; assay.
XX
OS Synthetic.
XX
PN JP03190898-A.
XX
PD 20-AUG-1991.
XX
PF 21-DEC-1989; 89JP-00329746.
XX
PR 21-DEC-1989; 89JP-00329746.
XX
PA (SHIM-) YG SHIMA KENKYUSHO.
XX
DR WPI; 1991-286117/39.
XX
PT Synthetic polypeptide reactive with HCV antibody - used to prepare a
PT solid layer reagent for measurement of HCV antibodies - having antigen-
PT determining gp. common to HCV and HCV antibody-measuring agent contg.
XX
PS solid layer for polypeptide fixation.
XX
SQ Claim 2; Page 1; 8pp; Japanese.
XX
CC This peptide has an epitope common to HCV. The peptide is strongly
CC reactive with anti-HCV antibodies from HCV-infected patients. By fixing
CC the peptide to a solid layer, e.g. a microtitre plate, it can be used in
CC an enzyme-linked assay to detect anti-HCV Abs
XX
SQ Sequence 42 AA;
Query Match 6.8%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||||
|||||||

RESULT 631
AAR30683
ID AAR30683 standard; peptide; 42 AA.
XX
AC AAR30683;
XX
DT 25-MAR-2003 (revised)
DT 11-MAY-1993 (first entry)
XX
DE Non-structural HCV peptide No. 19.
XX
KW Hepatitis C virus; HCV; open reading frame; "common" sequence;
KW capsid protein; epitope; immunoassay; antibody; diagnosis; NANBH;
KW non-A, non-B hepatitis; competitive; inhibition assay.
XX
OS Hepatitis C virus.
XX
PN WO9222571-A1.
XX
PD 23-DEC-1992.
XX
PF 29-APR-1992; 92WO-US003635.
XX
SQ Sequence 42 AA;
Query Match 6.8%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||||
|||||||

RESULT 632
AAR33873
ID AAR33873 standard; peptide; 42 AA.
XX
AC AAR33873;
XX
DT 25-MAR-2003 (revised)
DT 19-JUL-1993 (first entry)
XX
DE Polypeptide p1694 comprising HCV viral antigen.
XX
KW Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG;
KW p408.
XX
OS Synthetic.
XX
PN WO9306247-A1.
XX
PD 01-APR-1993.
XX
PF 16-SEP-1992; 92WO-US007813.
XX
PR 16-SEP-1991; 91US-00760292.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Lesniewski RR, Leung TK;
XX
DR WPI; 1993-117563/14.
XX
PT Assay for detecting presence of antibody to hepatitis C viral antigen -
PT by contacting sample with polypeptide contg. at least one epitope of
PT virus antigen.
XX
PS Disclosure; Page 13; 63pp; English.
XX
CC The synthetic peptide p1694 represents amino acid residues 1694-1735 of
CC the hepatitis C viral antigen. The peptide may be used in an assay to
CC detect antibodies to HCV and thus to diagnose chronic HCV infection. See

```

```

PR 13-JUN-1991; 91US-00714471.
PR 20-JUN-1991; 91US-00718052.
XX
PA (BAXT ) BAXTER DIAGNOSTICS INC.
XX
PI Leahy DC, Todd JA, Jolley ME;
XX
DR WPI; 1993-018073/02.
XX
PT Synthetic peptide with sequence encoded by hepatitis-C virus -
PT for immunoassay for antigens for diagnosis of non-A, non-B hepatitis.
XX
PS Disclosure; Fig 1D; 66pp; English.
XX
CC The sequences given in AAR30665-89 represent fragments of the Hepatitis C
CC virus (HCV) amino acid sequence. They represent the beginning of the HCV
CC open reading frame to amino acid 38 and encompass the "common" sequence.
CC These peptides are contained in the capsid protein of the virus and
CC themselves contain epitope groups. These peptides can be used in
CC immunoassays for HCV antibodies in the diagnosis of non-A, non-B
CC hepatitis (NANBH), and in competitive inhibition assay for detecting HCV
CC specific antibodies. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 42 AA;
Query Match 6.8%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||||
|||||||

```

```

RESULT 632
AAR33873
ID AAR33873 standard; peptide; 42 AA.
XX
AC AAR33873;
XX
DT 25-MAR-2003 (revised)
DT 19-JUL-1993 (first entry)
XX
DE Polypeptide p1694 comprising HCV viral antigen.
XX
KW Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG;
KW p408.
XX
OS Synthetic.
XX
PN WO9306247-A1.
XX
PD 01-APR-1993.
XX
PF 16-SEP-1992; 92WO-US007813.
XX
PR 16-SEP-1991; 91US-00760292.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Lesniewski RR, Leung TK;
XX
DR WPI; 1993-117563/14.
XX
PT Assay for detecting presence of antibody to hepatitis C viral antigen -
PT by contacting sample with polypeptide contg. at least one epitope of
PT virus antigen.
XX
PS Disclosure; Page 13; 63pp; English.
XX
CC The synthetic peptide p1694 represents amino acid residues 1694-1735 of
CC the hepatitis C viral antigen. The peptide may be used in an assay to
CC detect antibodies to HCV and thus to diagnose chronic HCV infection. See

```

```
CC also AAR33861-87. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 42 AA;

Query Match      6.8%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||

RESULT 633
AAR67544
ID AAR67544 standard; peptide; 42 AA.
AC AAR67544;
XX
XX 22-AUG-1995 (first entry)
DT
XX Hepatitis C virus type I peptide antigen.
DE
XX
XX Hepatitis C virus type I; HCV; non-A non-B; peptide antigen;
KW identification and diagnosis.
XX
OS Hepatitis C virus.
XX
XX JP06317597-A.
PN
PD 15-NOV-1994.
XX
XX 30-APR-1993; 93JP-00104753.
XX
XX 30-APR-1993; 93JP-00104753.
XX (TOFU) TONEN CORP.
XX
XX WPI; 1995-033119/05.
DR
XX
XX Peptide antigen of hepatitis C virus of group 1 - used in rapid
PT identification test and for diagnosis of infection.
XX
XX Claim 1; Page 4-5; 5pp; Japanese.
FS
XX
XX AAR67544 is a hepatitis C virus (HCV) type I peptide antigen. It can be
CC used in an identification test to distinguish between the type I, and
CC type II HCV infection
XX
SQ Sequence 42 AA;

Query Match      6.8%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||

RESULT 634
ADT77852
ID ADT77852 standard; peptide; 42 AA.
XX
XX ADT77852;
AC
XX
XX 13-JAN-2005 (first entry)
DT
XX
XX Hepatitis C virus non-structural protein NS4 peptide.
DE
XX
XX HCV; non-structural protein; NS4; adjuvant; antiinflammatory;
KW immunosuppressive; antimicrobial; cytostatic; antiasthmatic;
KW antiallergic; neuroprotective; antidiabetic; antirheumatic;
```

```
KW antiarthritic; dermatological; ophthalmological.
XX
XX Hepatitis C virus.
XX
XX WO2004089978-A2.
XX
XX 21-OCT-2004.
XX
XX 08-APR-2004; 2004WO-1E000054.
XX
XX 11-APR-2003; 2003IE-00000279.
XX
XX (QUEE-) QUEEN ELIZABETH COLLEGE DUBLIN.
XX
XX Mills KHG, Brady MT;
XX
XX WPI; 2004-748721/73.
XX
XX Therapeutic composition for preventing or treating inflammatory or immune
XX -mediated disorders comprises a Hepatitis C virus (HCV) protein activated
XX by an agent that suppresses inflammatory cytokine or promotes IL-10
XX production.
XX
XX Disclosure; SEQ ID NO 2; 43pp; English.
XX
XX The present sequence is the protein sequence of hepatitis C virus (HCV)
XX genotype 1a non-structural protein NS4 peptide fragment comprising amino
XX acid residues 1694-1735 of the HCV polyprotein. The invention provides a
XX method for the treatment and/or prophylaxis of an inflammatory and/or
XX immune-mediated disorder and/or disorders associated with transplantation
XX comprising the step of administering an HCV protein (NS4 or NS3) or its
XX derivative, mutant, fragment or variant, or the product of cells
XX activated thereby. In one case, the HCV protein or its product suppresses
XX inflammatory cytokine production and also promotes interleukin-10
XX production, particularly by peripheral blood mononuclear cells and/or
XX monocytes. It may also inhibit dendritic cell activation, inhibit the
XX induction or activation of Th1 or Th2 cells, modulate inflammatory cytokine
XX ligand-induced NFkappaB activation, or modulate inflammatory cytokine
XX production induced by infection or trauma. A claimed therapeutic
XX composition and a claimed vaccine adjuvant comprise HCV NS4 or its
XX derivative, mutant, fragment or variant or the product of cells activated
XX thereby. It can be used in the treatment or prophylaxis of sepsis or
XX acute inflammation induced by infection, trauma or injury, chronic
XX inflammatory disease, graft rejection or graft versus host disease, an
XX immune mediated disease involving Th1 responses, a disease or condition
XX involving toll-like receptor dependent signalling, an immune mediated
XX disease involving inflammatory cytokines including tumour necrosis factor
XX -alpha and interleukin-1, Crohn's disease, inflammatory bowel disease,
XX multiple sclerosis, type 1 diabetes, rheumatoid arthritis, systemic lupus
XX erythematosus, uveitis, allergy or asthma (all claimed). The present 42-
XX amino acid NS4 peptide fragment from HCV genotype 1a was shown to retain
XX the immunomodulatory activity observed in the full-length NS4 protein
XX ADT77851 from HCV genotype 1b.
XX
XX Sequence 42 AA;

Query Match      6.8%; Score 8; DB 8; Length 42;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||

RESULT 635
AAR13355
ID AAR13355 standard; protein; 43 AA.
XX
XX AAR13355;
AC
XX
XX 25-MAR-2003 (revised)
DT
XX 23-OCT-1991 (first entry)
```

XX P1694 HCV antigen (1694-1735).
 XX C100-3; hepatitis C virus; immunoassay; epitope.
 XX Synthetic.
 XX AU9068390-A.
 XX 27-JUN-1991.
 XX 21-DEC-1990; 90AU-00068390.
 XX 22-DEC-1989; 89US-00456162.
 XX 07-NOV-1990; 90US-00610180.
 XX (ABBO) ABBOTT LAB.
 XX (LESN/) LESNIEWSKI R R.
 XX WPI; 1991-238393/33.
 XX Immunological assays for hepatitis C virus antibody - by using
 XX polypeptide(s) contg. epitope(s) of hepatitis C virus antigens.
 XX Claim 1; Page 45; 62pp; English.
 XX The polypeptide may be prepared by solid phase synthesis fragment
 XX coupling (pref.) or using recombinant technology. The assay has increased
 XX sensitivity and is more specific than assays using the polypeptide C100-3
 XX (EP-318216). See also AAO13146-48 and AAR13343-65. (Updated on 25-MAR-
 XX 2003 to correct PA field.)
 XX Sequence 43 AA;
 Query Match 6.8%; Score 8; DB 2; Length 43;
 Best Local Similarity 100.0%; Pred. No. 9.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 60 DEMEECSQ 67
 Db 14 DEMEECSQ 21
 RESULT 636
 AAR74228
 ID AAR74228 standard; peptide; 46 AA.
 AC AAR74228;
 XX 25-MAR-2003 (revised)
 DT 24-DEC-1995 (first entry)
 XX HCV NS4 peptide SSAL.
 XX Structured antigenic peptide library; SSAL; vaccine; diagnostic;
 KW therapeutic; hepatitis C virus; HCV.
 XX Synthetic.
 XX Key Location/Qualifiers
 FH Misc-difference 1 /note= "S8;N2"
 FT Misc-difference 2 /note= "G8;D1;Q1"
 FT Misc-difference 3 /note= "K6;R4"
 FT Misc-difference 4 /note= "P8;V1"
 FT Misc-difference 5 /note= "A8;V2"
 FT Misc-difference 6 /note= "I6;V4"
 FT Misc-difference 7

FT Misc-difference 10 /note= "I6;V2;A2"
 FT Misc-difference 12 /note= "R8;K2"
 FT Misc-difference 15 /note= "V9;I1"
 FT Misc-difference 16 /note= "R3;Q5;E2"
 FT Misc-difference 24 /note= "E8;A2"
 FT Misc-difference 25 /note= "S4;A6"
 FT Misc-difference 26 /note= "Q4;S6"
 FT Misc-difference 27 /note= "H8;K1;R1"
 FT Misc-difference 28 /note= "L8;A2"
 FT Misc-difference 29 /note= "P8;A2"
 FT Misc-difference 30 /note= "Y8;L2"
 FT Misc-difference 32 /note= "I9;F1"
 FT Misc-difference 34 /note= "Q8;E2"
 FT Misc-difference 35 /note= "M8;Q2"
 FT Misc-difference 36 /note= "M4;Q4;R2"
 FT Misc-difference 39 /note= "L8;M1;I1"
 FT Misc-difference 40 /note= "Q8;M2"
 FT Misc-difference 42 /note= "F8;L2"
 FT Misc-difference 44 /note= "Q8;S2"
 FT Misc-difference 45 /note= "A8;I2"
 FT Misc-difference 45 /note= "L8;Q2"
 XX WO9511998-A1.
 XX 04-MAY-1995.
 XX 26-OCT-1994; 94WO-US012268.
 XX 26-OCT-1993; 93US-00143412.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY, Zamb TJ, Ye J, Kaminsky SM, Hosein B, Nixon DF;
 XX Koif CW, Kowalski J, Walfield AM;
 XX WPI; 1995-178890/23.
 XX Structured antigenic peptide libraries contain some invariant amino acids
 XX - accommodate variations in antigenic structure so are effective against
 XX many different strains of e.g. rapidly mutating viruses in vaccines.
 XX Claim 4; Page 69; 216pp; English.
 XX In a structured synthetic antigen library, specific amino acids and their
 XX frequency of appearance at a variant locus within aligned peptide
 XX sequences are defined by the primary sequences of the several variants
 XX that make up the alignment used to construct the antigen peptide library.
 XX The SSAL given in AAR74228 illustrates a primary amino acid sequence of
 XX antigenic HCV NS4 peptide from 8 HCV strains. It was used to detect HCV
 XX antibodies in patients from geographically distinct regions and to
 XX develop a sensitive, specific immunoassay for HCV. (Updated on 25-MAR-
 XX 2003 to correct PI field.)

```

SQ Sequence 46 AA;
Query Match      6.8%; Score 8; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 17 DEMEECSQ 24

RESULT 637
AAR20771
ID AAR20771 standard; protein; 47 AA.
AC AAR20771;
XX
XX 25-MAR-2003 (revised)
DT 05-MAY-1992 (first entry)
XX
XX Peptide IIH based on immunoreactive region of Hepatitis C virus.
DE
XX Non-A, non-B hepatitis virus; non-structural protein; vaccine.
XX
XX Synthetic.
XX
XX EP468527-A.
XX
XX 29-JAN-1992.
XX
XX 26-JUL-1990; 90US-00558799.
XX
XX 26-JUL-1990; 90US-00558799.
PR 07-FEB-1991; 91US-00651735.
PR 11-MAR-1991; 91US-00667275.
PR 24-JUN-1991; 91US-00719819.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Chang YW, Hosein B;
XX
XX WPI; 1992-034279/05.
XX
XX New synthetic peptide specific for HCV antibodies - for detection of HCV
or NANBHV e-g. by enzyme-linked immunosorbent assay and is immunogen for
preparation of vaccines.
XX
XX Claim 7; Page 92; 98pp; English.
XX
XX Peptide IIH is from the non-structural protein region of HCV. It is used
in a composition with peptide 11 (see AAR20761), peptide VIII (see
AAR20770) and optionally with peptide 8 (AAR20758) and/or 12 (AAR20762)
for the detection of antibodies to HCV and diagnosis of NANBH. The C-
terminal amino acid of peptide IIH may be amidated. See AAR20751-R20782.
(Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 47 AA;
Query Match      6.8%; Score 8; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 18 DEMEECSQ 25

RESULT 638
AAR69544
ID AAR69544 standard; peptide; 47 AA.
AC AAR69544;
XX
XX 25-MAR-2003 (revised)
DT 03-SEP-1995 (first entry)
XX
XX Anti-HCV antibody immunoreactive peptide II H.
XX
XX Hepatitis C virus; HCV; non-A non-B; antibodies; vaccines;
immunoreactive peptide II H; infection detection assay.
XX
XX Synthetic.
XX
XX WO9500670-A1.
XX
XX 05-JAN-1995.
XX
XX 22-JUN-1994; 94WO-US007088.
XX
XX 28-JUN-1993; 93US-00083947.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY, Hosein B;
XX
XX WPI; 1995-052105/07.
XX
XX Linear and branched peptide(s) comprising, e.g. PepB and PepC - useful in
vaccines and assays for non-A, non-B hepatitis.
XX
XX Claim 14; Page 48; 58pp; English.
XX
XX AAR69523-R69545 are anti-hepatitis C virus (HCV) antibody immunoreactive
peptides, based on the PepB (AAR69529) or the PepC (AAR69534) sequences.
These peptides, in linear or branched dimer compns. are used in
CC immunoassays for the detection of HCV infections, the peptide compns.
CC may also be useful in vaccines against these infections. (Updated on 25-
MAR-2003 to correct PN field.)
XX
XX Sequence 47 AA;
Query Match      6.8%; Score 8; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 18 DEMEECSQ 25

RESULT 639
AAY15424
ID AAY15424 standard; peptide; 47 AA.
XX
XX AAY15424;
AC
XX
XX 26-JUL-1999 (first entry)
DT
XX
XX Prototype peptide which is reactive with HCV NS-4 antibodies.
XX
XX Linear peptide; branched peptide; immunoreactive; HCV antibody; NS-3;
NS-4; NS-5; CP; immunoassay reagent; HCV detection; diagnosis; infection.
XX
XX Hepatitis C virus.
OS
XX GB2294690-A.
PN
XX 08-MAY-1996.
PD
XX
XX 19-DEC-1994; 94GB-00025604.
XX
XX 01-NOV-1994; 94US-00333573.
PR
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY, Hosein BH;
PI

```


XX WPI; 1989-159274/22.
DR N-PSDB; AAN92073.
XX
XX Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
XX
XX Claim 15; Fig 1; 139pp; English.
PS
XX It is the putative amino acid sequence of a polypeptide encoded in
CC hepatitis C virus (HCV) cDNA insert in clone 5-1-1. It is an epitope to
CC which antibodies in sera from humans with non-A, non-B hepatitis bind.
CC Portions of it could be used as immunoassay reagents and vaccines and to
CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
CC infection. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-
CC MAR-2003 to correct PI field.)
CC
CC Revised record issued on 09-SEP-2004 : Correction to DE line
CC
XX Sequence 51 AA;
SQ
Query Match 6.8%; Score 8; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 22 DEMEECSQ 29
|||||
RESULT 642
AAP90134
ID AAP90134 standard; protein; 51 AA.
XX
AC AAP90134;
XX
XX 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
XX Sequence of hepatitis C virus cDNA insert in clone 5-1-1.
DE
XX Hepatitis C virus; clone 5-1-1; probe; vaccine.
XX
XX Pan troglodytes.
XX
XX GB2212511-A.
FN
XX 26-JUL-1989.
PD
XX 18-NOV-1988; 88GB-00027024.
PF
XX 18-NOV-1987; 87US-00122714.
PR
XX 30-DEC-1987; 87US-00139886.
PR
XX 26-FEB-1988; 88US-00161072.
PR
XX 26-OCT-1988; 88US-00263584.
XX
XX (CHIR) CHIRON CORP.
FA
XX
XX Houghton M, Choo QL, Kuo G;
PI
XX WPI; 1989-215054/30.
DR
DR N-PSDB; AAN90303.
XX
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
XX Disclosure; Fig 1; 30pp; English.
PS
XX The sequence is the peptide sequence of the hepatitis C virus (HCV) cDNA
CC insert in clone 5-1-1 (see AAN90303). The peptides react with antibodies
CC present in patients with non-A non-B hepatitis (NANBH). The polypeptides
CC are used to diagnose HCV-induced NANBH, to raise antibodies for

CC immunoassay or treatment, or to produce vaccines. See also AAN90304-36.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
XX Sequence 51 AA;
SQ
Query Match 6.8%; Score 8; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 22 DEMEECSQ 29
|||||
RESULT 643
ABB77222
ID ABB77222 standard; protein; 54 AA.
XX
AC ABB77222;
XX
XX 28-JUN-2002 (first entry)
DT
XX Selected interacting domain (SID) polypeptide 8.
DE
XX SID; selected interacting domain; HCV; hepatitis C virus; liver disease;
KW liver cancer; virucide; hepatotropic; antiinflammatory; antibacterial.
XX
XX Hepatitis C virus strain H77.
OS
XX
XX EP1178116-A1.
PN
XX 06-FEB-2002.
PD
XX 03-AUG-2000; 2000EP-00402225.
PF
XX 03-AUG-2000; 2000EP-00402225.
PR
XX (HYBR-) HYBRIGENICS SA.
PA
XX
XX Legrain P, Whiteside S, Wojcik J;
PI
XX WPI; 2002-208115/27.
DR
DR N-PSDB; ABL55554.
XX
XX New selected interacting domain polypeptides and polynucleotides, useful
PT for treating or preventing infections or pathologies caused by hepatitis
PT C virus (HCV) or those linked to HCV infection.
XX
XX Claim 7; SEQ ID NO 8; 61pp + Sequence Listing; English.
PS
XX The invention relates to nucleic acids encoding polypeptides which are
CC termed SID polypeptides (selected interacting domain). These polypeptides
CC are the final products of a double selection method involving a first
CC step of selection of Hepatitis C virus (HCV)-derived polynucleotides
CC through a two-hybrid system, and a second selection step involving an
CC alignment between the different polynucleotides selected at the first
CC step. The activity of polypeptides of the invention may be described as,
CC virucide, hepatotropic, antiinflammatory and antibacterial. The
CC polypeptide, polynucleotide and compositions comprising them are useful
CC for treating or preventing viral or a bacterial infection, specifically
CC infections or pathologies caused by HCV, or those pathologies linked to
CC HCV infection. These may include liver disease and liver cancer. The
CC current sequence represents a selected interacting domain (SID)
CC polypeptide. Note: The sequence data for this patent is not represented
CC in the specification, but is based on sequence information supplied by
CC the European Patent Office
XX
XX Sequence 54 AA;
SQ
Query Match 6.8%; Score 8; DB 5; Length 54;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMERCSQ 67
 DB 28 DEMERCSQ 35
 |||||
 |||||

RESULT 644
 AAR26489
 ID AAR26489 standard; protein; 55 AA.
 AC AAR26489;
 XX
 DT 10-NOV-1992 (first entry)
 XX
 DE Fragment of open reading frame of hepatitis C virus C-100-3 construct.
 XX
 KW Hepatitis C virus; C-100-3 construct; peptide fragment; immunoassay;
 KW reagent; probes; antibodies; HIV; HTLV; hepatitis B virus;
 KW Treponema pallidum; open reading frame.
 XX
 OS Hepatitis C virus.
 XX

Key Location/Qualifiers
 FH Peptide 1..40
 FT /note= "claim 3, page 72 (peptide 4074)"
 FT Peptide 1..19
 FT /note= "claim 3, page 74 (peptide 4060)"
 FT Peptide 1..15
 FT /note= "claim 3, page 73 (peptide 4056)"
 FT Peptide 8..40
 FT /note= "claim 3, page 72 (peptide 4073)"
 FT Peptide 9..23
 FT /note= "claim 3, page 73 (peptide 4055)"
 FT Peptide 16..45
 FT /note= "claim 3, page 73 (peptide 4090)"
 FT Peptide 16..40
 FT /note= "claim 3, page 73 (peptide 4072)"
 FT Peptide 17..31
 FT /note= "claim 3, page 73 (peptide 4054)"
 FT Peptide 24..55
 FT /note= "claim 3, page 73 (peptide 4082)"
 FT Peptide 24..40
 FT /note= "claim 3, page 73 (peptide 4071)"
 FT Peptide 25..39
 FT /note= "claim 3, page 73 (peptide 4053)"
 FT Peptide 33..47
 FT /note= "claim 3, page 74 (peptide 4052)"
 FT Peptide 41..55
 FT /note= "claim 3, page 73 (peptide 4081)"

EP484787-A.
 13-MAY-1992.
 28-OCT-1991; 91EP-00118349.
 03-NOV-1990; 90DE-04034982.
 19-APR-1991; 91DE-04112743.
 19-JUN-1991; 91DE-04120281.
 28-JUN-1991; 91DE-04121431.
 (BEHM) BEHRINGER AG.
 Krupka U, Stuber W, Gerken M, Brust S;
 WPI; 1992-160606/20.
 Peptide(s) for hepatitis-C virus detection - their use in immunological tests, and single test for the detection of different epitope(s) from different pathogens.
 Claim 1; Page 71; 94pp; German.
 The peptide is residues 121-175 of the open reading frame of hepatitis C

CC virus construct C-100-3 (this sequence is a fragment of that given in EP-318216). The fragments listed in the features table are also claimed. The fragments react specifically with antibodies to hepatitis C virus. They are useful as assay reagents for diagnosis of HCV infection, allowing a decrease in false positives cf. known methods. The DNA sequences encoding the peptides are useful as hybridisation probes for detection and/or determin. of HCV. Antibodies raised to the peptide are useful CC diagnostically and therapeutically, esp. as immunoassay reagents. By using 2 or more peptides with specific affinity for early and late anti-HCV antibodies, differential diagnosis of acute and chronic infections CC may be achieved of HIV, HTLV, hepatitis B virus and Treponema pallidum XX

SQ Sequence 55 AA;
 Query Match 6.8%; Score 8; DB 2; Length 55;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMERCSQ 67
 DB 18 DEMERCSQ 25
 |||||
 |||||

RESULT 645
 AAR26412
 ID AAR26412 standard; peptide; 55 AA.
 XX
 AC AAR26412;
 XX
 DT 24-OCT-2003 (revised)
 DT 19-NOV-1992 (first entry)
 XX
 DE HCV C-100-3 peptide 121-175.
 XX
 KW Immunoassay; vaccine; early; late; diagnosis.
 XX
 OS Hepatitis C virus; C-100-3.
 XX

Key Location/Qualifiers
 FT Peptide 1..40
 FT /note= "claim 3; page 18-20"
 FT Peptide 1..19
 FT /note= "claim 3; page 18-20"
 FT Peptide 1..15
 FT /note= "claim 3; page 18-20"
 FT Peptide 8..40
 FT /note= "claim 3; page 18-20"
 FT Peptide 9..23
 FT /note= "claim 3; page 18-20"
 FT Peptide 15..44
 FT /note= "claim 3; page 18-20"
 FT Peptide 16..40
 FT /note= "claim 3; page 18-20"
 FT Peptide 17..31
 FT /note= "claim 3; page 18-20"
 FT Peptide 24..55
 FT /note= "claim 3; page 18-20"
 FT Peptide 24..40
 FT /note= "claim 3; page 18-20"
 FT Peptide 25..39
 FT /note= "claim 3; page 18-20"
 FT Peptide 33..47
 FT /note= "claim 3; page 18-20"
 FT Peptide 41..55
 FT /note= "claim 3; page 18-20"

DE4034982-A.
 07-MAY-1992.
 03-NOV-1990; 90DE-04034982.
 03-NOV-1990; 90DE-04034982.

```

XX PA (BEHW ) BEHRINGERWERKE AG.
XX PI Krupka U, Stuber W;
XX XX
XX DR WPI; 1992-160123/20.
XX XX
XX PT New peptide(s) corresponding to part of HCV C-100-3 construct - useful in
XX PT vaccines and in binding assays, e.g. to distinguish between acute and
XX PT chronic infection.
XX PS Claim 2; Page 18; 22pp; German.
XX XX
XX CC The peptides comprising amino acids 121-175 (AAR26412) and 337- 363
XX CC (AAR26413), as well as the fragments indicated in the features table,
XX CC correspond to portions of the HCV C-100-3 construct (see WO8904669). The
XX CC peptides are useful as immunoassay reagents for diagnosis of HCV
XX CC infections, and for the prodn. of vaccines against HCV infections. By
XX CC using two or more peptides with specific affinity for early and late anti
XX CC -HCV antibodies, differential diagnosis of acute and chronic infections
XX CC may be achieved. (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 55 AA;
XX
XX Query Match 6.8%; Score 8; DB 2; Length 55;
XX Best Local Similarity 100.0%; Pred. No. 11;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 60 DEMEECSQ 67
XX Db 18 DEMEECSQ 25
XX
XX RESULT 646
XX AAR13345
XX ID AAR13345 standard; protein; 67 AA.
XX XX
XX AC AAR13345;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 23-OCT-1991 (first entry)
XX XX
XX DE P1684 HCV antigen (1684-1750).
XX XX
XX KW C100-3; hepatitis C virus; immunoassay; epitope.
XX OS Synthetic.
XX XX
XX PN AU9068390-A.
XX XX
XX PD 27-JUN-1991.
XX XX
XX PF 21-DEC-1990; 90AU-00068390.
XX XX
XX PR 22-DEC-1989; 89US-00456162.
XX PR 07-NOV-1990; 90US-00610180.
XX XX
XX PA (ABBO ) ABBOTT LAB.
XX PA (LESN/) LESNIEWSKI R R.
XX XX
XX DR WPI; 1991-238393/33.
XX DR N-PSDB; AAQ13148.
XX XX
XX PT Immunological assays for hepatitis C virus antibody - by using
XX PT polypeptide(s) contg. epitope(s) of hepatitis C virus antigens.
XX XX
XX PS Claim 1; Page 45; 62pp; English.
XX XX
XX CC The polypeptide may be prepared by solid phase synthesis fragment
XX CC coupling or using recombinant technology. The assay has increased
XX CC sensitivity and is more specific than assays using the polypeptide C100-3
XX CC (EP-318216). See also AAQ13146-48 and AAR13343-65. (Updated on 25-MAR-
XX CC 2003 to correct PA field.)

```

```

XX SQ Sequence 67 AA;
XX
XX Query Match 6.8%; Score 8; DB 2; Length 67;
XX Best Local Similarity 100.0%; Pred. No. 13;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 60 DEMEECSQ 67
XX Db 23 DEMEECSQ 30
XX
XX RESULT 647
XX AAR33871
XX ID AAR33871 standard; peptide; 67 AA.
XX XX
XX AC AAR33871;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 19-JUL-1993 (first entry)
XX XX
XX DE Polypeptide p1684 comprising HCV viral antigen.
XX XX
XX KW Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG;
XX KW p408.
XX OS Synthetic.
XX XX
XX PN WO9306247-A1.
XX PN 01-APR-1993.
XX XX
XX PF 16-SEP-1992; 92WO-US007813.
XX XX
XX PR 16-SEP-1991; 91US-00760292.
XX XX
XX PA (ABBO ) ABBOTT LAB.
XX XX
XX PI Lesniewski RR, Leung TK;
XX DR WPI; 1993-117563/14.
XX XX
XX PT Assay for detecting presence of antibody to hepatitis C viral antigen -
XX PT by contacting sample with polypeptide contg. at least one epitope of
XX PT virus antigen.
XX PS Disclosure; Page 13; 63pp; English.
XX XX
XX CC The synthetic peptide p1684 represents amino acid residues 1684-1750 of
XX CC the hepatitis C viral antigen. The peptide may be used in an assay to
XX CC detect antibodies to HCV and thus to diagnose chronic HCV infection. See
XX CC also AAR33861-87. (Updated on 25-MAR-2003 to correct PN field.)
XX SQ Sequence 67 AA;
XX
XX Query Match 6.8%; Score 8; DB 2; Length 67;
XX Best Local Similarity 100.0%; Pred. No. 13;
XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 60 DEMEECSQ 67
XX Db 23 DEMEECSQ 30
XX
XX RESULT 648
XX AAR51389
XX ID AAR51389 standard; peptide; 80 AA.
XX XX
XX AC AAR51389;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 26-OCT-1994 (first entry)
XX XX

```

DE Branched peptide H1.
 XX Branch; lysine dimer; lysine octamer; immunoassay; reagent;
 KW Non-A, non-B hepatitis; NANBH; NANBH-related antibody; hepatitis C; HCV;
 KW ELISA; passive haemagglutination; vaccine.
 XX Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 80
 FT /note= "Linked to a lysine dimer to form branched
 FT molecule"
 XX
 PN WO9406826-A1.
 XX
 PD 31-MAR-1994.
 XX
 XX 15-SEP-1993; 93WO-US008638.
 PF
 XX 15-SEP-1992; 92US-00946054.
 PR
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA
 XX Wang CY, Hosein B;
 PI
 XX WPI; 1994-118396/14.
 DR
 XX New branched peptides for detecting non-A, non-B hepatitis antibodies -
 PT to identify carriers and diagnose hepatitis C infection, contain at least
 PT one specific epitope.
 XX
 PS Claim 18; Page 12; 43pp; English.
 XX
 CC The sequences given in AAR51389-426 are peptides which form branched
 CC peptides through either a lysine dimer or a lysine octamer at the C-
 CC terminal. These branched peptides are immunoassay reagents for detecting
 CC Non-A, non-B hepatitis (NANBH)-related antibodies. They may be used in
 CC the diagnosis of hepatitis C (HCV) and for the identification of
 CC carriers, particularly by ELISA or passive haemagglutination. They can
 CC also be used as vaccines to protect against HCV infection. (Updated on 25
 CC -MAR-2003 to correct PN field.)
 XX
 SQ Sequence 80 AA;
 Query Match 6.8%; Score 8; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMEECSQ 67
 Db 40 DEMEECSQ 47
 |||||
 RESULT 649
 AAR13557
 ID AAR13557 standard; peptide; 90 AA.
 XX
 AC AAR13557;
 XX
 XX 25-MAR-2003 (revised)
 DT 28-OCT-1991 (first entry)
 XX
 XX HCV C-100 protein immunodominant region.
 DE
 XX Hepatitis C virus; non-A non-B hepatitis virus; diagnosis; C-100 protein;
 KW core protein; vaccines; NANBH.
 XX
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..40
 FT /label= IIIF
 FT Peptide 11..50

FT Peptide /label= IV
 FT 16..57
 FT /label= IIG
 FT 31..70
 FT /label= IE
 FT 41..81
 FT /label= VIE
 FT 51..90
 FT /label= VE
 XX
 PN EP442394-A.
 XX
 PD 21-AUG-1991.
 XX
 XX 08-FEB-1991; 91EP-00101787.
 PF
 XX 16-FEB-1990; 90US-00481348.
 PR
 XX 16-APR-1990; 90US-00510153.
 PR
 XX 26-JUL-1990; 90US-00558799.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 XX WPI; 1991-247104/34.
 DR
 XX
 XX New synthetic peptide(s) from immuno-dominant regions of virus - for
 PT diagnosis of hepatitis C virus and non-A, -B hepatitis infection, esp.
 PT using enzyme-linked immuno-sorbent assay.
 XX
 PS Disclosure; Page 15; 93pp; English.
 XX
 CC In selecting regions of the HCV protein for epitope analysis, peptides in
 CC the 40 mer size range with amino acid sequences covering the complete HCV
 CC C-100 protein and the core protein were synthesised. These were tested
 CC for their reactivity with serum from a patient positively diagnosed with
 CC HCV infection. The indicated six overlapping peptides from the HCV C-100
 CC protein region were found to have specific immunoreactivity with the
 CC positive control serum. The peptides may be used in highly sensitive and
 CC accurate methods for the early detection of antibodies to HCV in body
 CC fluids and the diagnosis of NANBHV infection. Because of their high
 CC immunoreactivity, the peptides are also useful in stimulating prodn. of
 CC antibodies to HCV and in vaccines to prevent HCV or NANBHV infection. See
 CC also AAR13558 for core protein immunodominant peptides. (Updated on 25-
 CC MAR-2003 to correct PA field.)
 XX
 SQ Sequence 90 AA;
 Query Match 6.8%; Score 8; DB 2; Length 90;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMEECSQ 67
 Db 28 DEMEECSQ 35
 |||||
 RESULT 650
 AAP92019
 ID AAP92019 standard; protein; 117 AA.
 XX
 AC AAP92019;
 XX
 XX 09-SEP-2004 (revised)
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX
 XX Polypeptide encoded in the HCV cDNA insert in clone 81.
 DE
 XX Hepatitis C virus (HCV); non-A, non-B hepatitis (NANBH).
 KW
 XX Hepatitis C virus.
 OS
 XX Unidentified.

XX EP318216-A.
PN
XX
XX 31-MAY-1989.
XX
XX 18-NOV-1988; 88EP-00310922.
XX
XX 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
XX (CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
XX
XX WPI; 1989-159274/22.
DR N-PSDB; AAN92075.
XX
XX Purified hepatitis C virus - and associated nucleic acids and polypeptide(s).
XX
XX Claim 13; Fig 4; 139pp; English.
XX
XX It is the putative sequence encoded in the open reading frame of hepatitis C virus (HCV) cDNA insert in clone 81. It is an epitope, portions of which could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Revised record issued on 09-SEP-2004 : Correction to DE line
XX
XX Sequence 117 AA;
XX
Query Match 6.8%; Score 8; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 18 DEMEECSQ 25
RESULT 651
AAP90136
ID AAP90136 standard; protein; 117 AA.
XX
XX AAP90136;
AC
XX
XX 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
XX Sequence of hepatitis C virus cDNA insert in DNA 81.
DE
XX
XX Hepatitis C virus; DNA 81; probe; vaccine.
XX
XX Pan troglodytes.
OS
XX
XX GB2212511-A.
PN
XX
XX 26-JUL-1989.
XX
XX 18-NOV-1988; 88GB-00027024.
PF
XX
XX 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX

PA (CHIR) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
PI
XX
XX WPI; 1989-215054/30.
DR N-PSDB; AAN90305.
XX
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
PT
XX
XX Disclosure; Fig 4; 30pp; English.
PS
XX
XX The sequence is the peptide encoded by the hepatitis C virus (HCV) cDNA insert in DNA 81 (see AAN90305). The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
CC
XX
XX Sequence 117 AA;
SQ
Query Match 6.8%; Score 8; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 18 DEMEECSQ 25
RESULT 652
AAR33872
ID AAR33872 standard; peptide; 117 AA.
XX
XX AAR33872;
AC
XX
XX 25-MAR-2003 (revised)
DT 19-JUL-1993 (first entry)
XX
XX Polypeptide p1689 comprising HCV viral antigen.
DE
XX
XX Hepatitis C virus; NANBH; assay; antibody; p380-JH1; p380-J; p380LG; p408.
KW
XX
XX Synthetic.
OS
XX
XX WO9306247-A1.
PN
XX
XX 01-APR-1993.
PD
XX
XX 16-SEP-1992; 92WO-US007813.
PF
XX
XX 16-SEP-1991; 91US-00760292.
PR
XX
XX (ABBO) ABBOTT LAB.
PA
XX
XX Lesniewski RR, Leung TK;
PI
XX
XX WPI; 1993-117563/14.
DR
XX
XX Assay for detecting presence of antibody to hepatitis C viral antigen - by contacting sample with polypeptide contg. at least one epitope of virus antigen.
PT
XX
XX Disclosure; Page 13; 63pp; English.
PS
XX
XX The synthetic peptide p1689 represents amino acid residues 1689-1805 of the hepatitis C viral antigen. The peptide may be used in an assay to detect antibodies to HCV and thus to diagnose chronic HCV infection. See also AAR33861-87. (Updated on 25-MAR-2003 to correct PN field.)
CC
XX
XX Sequence 117 AA;
SQ
Query Match 6.8%; Score 8; DB 2; Length 117;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 18 DEMEECSQ 25

RESULT 653
AAAR13354
ID AAR13354 standard; protein; 118 AA.
XX
AC AAR13354;
XX
DT 25-MAR-2003 (revised)
DT 23-OCT-1991 (first entry)
DE P1689 HCV antigen (1689-1805).
XX
KW C100-3; hepatitis C virus; immunoassay; epitope.
XX
OS Synthetic.
XX
PN AU9068390-A.
XX
PD 27-JUN-1991.
XX
PF 21-DEC-1990; 90AU-00068390.
XX
PR 22-DEC-1989; 89US-00456162.
PR 07-NOV-1990; 90US-00610180.
XX
PA (ABBO) ABBOTT LAB.
PA (LESN/) LESNIEWSKI R R.
XX
DR WPI; 1991-238393/33.
XX
PT Immunological assays for hepatitis C virus antibody - by using
PT polypeptide(s) contg. epitope(s) of hepatitis C virus antigens.
XX
PS Claim 1; Page 45; 62pp; English.
XX
CC The polypeptide may be prepared by solid phase synthesis fragment
CC coupling (pref.) or using recombinant technology. The assay has increased
CC sensitivity and is more specific than assays using the polypeptide C100-3
CC (EP-318216). See also AAQ13146-48 and AAR13343-65. (Updated on 25-MAR-
CC 2003 to correct PA field.)
XX
SQ Sequence 118 AA;

Query Match 6.8%; Score 8; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 19 DEMEECSQ 26

RESULT 654
AAP92018
ID AAP92018 standard; protein; 128 AA.
XX
AC AAP92018;
XX
DT 09-SEP-2004 (revised)
DT 25-MAR-2003 (revised)
DT 02-MAR-1990 (first entry)
XX
DE Polypeptide encoded in a composite sequence of the HCV cDNA.
XX
KW Hepatitis C virus (HCV); non-A, non-B hepatitis (NANBH).
XX

OS Hepatitis C virus.
OS Unidentified.
XX
PN EP318216-A.
XX
PD 31-MAY-1989.
XX
PF 18-NOV-1988; 88EP-00310922.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX
PA (CHIR) CHIRON CORP.
PA (CHIR) CHIRON CORP.
XX
PI Houghton M, Choo QL, Kuo G;
XX
DR WPI; 1989-159274/22.
DR N-PSDB; AAN92074.
XX
PT Purified hepatitis C virus - and associated nucleic acids and
PT polypeptide(s).
XX
PS Claim 13; Fig 3; 139pp; English.
XX
CC It is encoded in the open reading frame of a composite nucleotide
CC sequence derived from overlapping hepatitis C virus (HCV) cDNA in clones
CC 81, 1-2, and 91, isolated using a synthetic sequence equivalent to a
CC fragment of HCV cDNA in clone 5-1-1. It is an epitope, portions of which
CC could be used as immunoassay reagents and vaccines and to generate
CC antibodies useful in diagnosis and passive immunotherapy for HCV
CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
CC field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
CC Revised record issued on 09-SEP-2004 : Correction to DE line
XX
SQ Sequence 128 AA;

Query Match 6.8%; Score 8; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 29 DEMEECSQ 36

RESULT 655
AAP90135
ID AAP90135 standard; protein; 128 AA.
XX
AC AAP90135;
XX
DT 25-MAR-2003 (revised)
DT 01-NOV-1989 (first entry)
XX
DE Sequence of hepatitis C virus cDNA insert in clone 5-1-1, 81 91 and 1-2.
XX
KW Hepatitis C virus; clone 5-1-1, 81, 91, 1-2; probe; vaccine.
XX
OS Pan troglodytes.
XX
PN GB2212511-A.
XX
PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.

```

PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
XX
XX WPI; 1989-215054/30.
DR N-PSDB; AAN90304.
XX
XX Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
XX Disclosure; Fig 3; 30pp; English.
XX
XX The sequence is the peptide encoded by the hepatitis C virus (HCV) cDNA
CC insert in clone 5-1-1, 81, 91 and 1-2 (see AAN90304). The polypeptides
CC are used to diagnose HCV-induced NANBH, to raise antibodies for
CC immunoassay or treatment, or to produce vaccines. (Updated on 25-MAR-2003
CC to correct PR field.)
XX
XX Sequence 128 AA;
SQ
Query Match 6.8%; Score 8; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db |||||
29 DEMEECSQ 36
RESULT 656
AAR58716
ID AAR58716 standard; protein; 270 AA.
XX
XX AAR58716;
AC
XX 27-AUG-2003 (revised)
DT 01-FEB-1995 (first entry)
XX
XX Non-A non-B hepatitis virus hybrid DNA.
DE
XX NANBHV; core; envelope; epitope; virus; fusion protein;
KW blood transmittable; detection; diagnosis; antigen.
XX
XX Non-A.
OS non-B hepatitis virus.
XX
XX JP06056893-A.
PN
XX 01-MAR-1994.
PD
XX 03-AUG-1992; 92JP-00226478.
PF
XX 03-AUG-1992; 92JP-00226478.
PR (JAPG ) NIPPON ZEON KK.
PA
XX WPI; 1994-194102/24.
DR N-PSDB; AAR71676.
XX
XX Fused protein comprising an envelope protein deriv. from a blood
PT transmittable Non-A, Non-B hepatitis virus - useful for the detection of
PT hepatitis virus.
XX
XX Disclosure; Page 8-9; 11pp; Japanese.
PS
XX A fused protein comprising an antigen protein derived from a blood
CC transmittable NANBHV is prepd. as follows: (1) viral DNA from a patient
CC with the disease is isolated; (2) cDNA (core gene) is prepd.; (3) DNA is
CC

```

```

CC inserted into cloning vector; (4) core gene, not including amino acids in
CC a hydrophobic region, is cloned. The envelope gene is ligated downstream
CC to the core gene and a gene encoding another antigen is further ligated
CC to the C-terminal of the envelope gene to improve expression. A
CC recombinant vector contg. the gene encoding the fusion protein is used
CC for the transformation of E.coli. The prod. expressed by E.coli is
CC isolated. Mass prodn. of a fused protein having an envelope protein
CC derived from blood transmittable NANBHV is possible. The fused protein
CC has an epitope of core protein, thus can be used for the detection and
CC diagnosis of NANBHV. (Updated on 27-AUG-2003 to correct OS field.)
XX
XX Sequence 270 AA;
SQ
Query Match 6.8%; Score 8; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db |||||
250 DEMEECSQ 257
RESULT 657
ADR46154
ID ADR46154 standard; peptide; 342 AA.
XX
XX ADR46154;
AC
XX 18-NOV-2004 (first entry)
DT
XX Brucella melitensis proline racemase.
DE
XX Proline racemase; enzyme; protozoacide; vaccine.
KW
XX Brucella melitensis.
OS Synthetic.
OS
XX Key Location/Qualifiers
FH Misc-difference 10..25
FT /note= "Amino acids of Motif I, illegible in Fig 6"
FT Misc-difference 143..146
FT /note= "Amino acids of Motif II, illegible in Fig 6"
FT Misc-difference 161..173
FT /note= "Amino acids of motif II, illegible in Fig 6"
FT Misc-difference 251..264
FT /note= "Amino acids of Motif III illegible in Fig 6"
XX
XX WO2004072223-A2.
PN
XX 26-AUG-2004.
PD
XX 11-FEB-2004; 2004WO-IB000861.
PF
XX 11-FEB-2003; 2003US-0446263P.
PR (INSP ) INST PASTEUR.
XX
XX Minoprio P, Chamond N, Degraeve W, Berneman A;
PI WPI; 2004-625840/60.
XX
XX New purified nucleic acid molecule encoding a racemase peptide isolated
PT from parasitic protozoan Trypanosoma cruzi, useful for preventing and/or
PT treating Chagas' disease.
XX
XX Example 12; Fig 6; 109pp; English.
PS
XX The present sequence is that of proline racemase from Brucella
CC melitensis. Sequence alignment of proline racemase sequences from
CC multiple organisms ADR46139-ADR46160 identified 4 motifs, designated
CC motif I ADR46161, motif II ADR46162, motif III ADR46120 and motif III*
CC ADR46121 that can be used to identify putative proline racemases in other
CC organisms. The invention relates to the identification and

```

CC characterisation of racemases, especially proline racemase, and
 CC definition of protein signatures of these racemases. It also relates to
 CC the identification of nucleic acid molecules encoding a peptide
 CC consisting of a motif characteristic of the protein signatures, to
 CC polypeptides consisting of these motifs, and to antibodies that bind such
 CC polypeptides. Methods are claimed for: producing a polypeptide consisting
 CC of a motif using bacterial, parasite or eukaryotic host cells; detecting
 CC a racemase; screening for an active molecule capable of inhibiting a
 CC racemase; detecting a D-amino acid, especially in a reaction medium
 CC comprising a sample from a patient afflicted with Alzheimer's disease,
 CC Parkinson's disease, renal disease or schizophrenia; detecting racemase
 CC activity in a reaction medium; and screening for molecules that modulate
 CC racemase activity, especially proline racemase from *Trypanosoma cruzi*
 CC (the causative agent of Chagas disease). An immunising composition
 CC containing a polypeptide comprising one of the motifs is also claimed.

XX Sequence 342 AA;

Query Match 6.8%; Score 8; DB 8; Length 342;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPAI 48
 Db 297 ELGGKPAI 304
 |||||

RESULT 658

ABM15881
 ID ABM15881 standard; protein; 399 AA.

XX AC ABM15881;

DT 26-SEP-2003 (first entry)

XX Mycobacterium tuberculosis mycobacterial antigen protein SEQ ID NO:127.

XX Mycobacterium tuberculosis; mycobacterial; antigen; infection; vaccine;
 KW tuberculostatic; mycobacterial peptide; mycobacterial infection.

XX Mycobacterium tuberculosis.

XX WO200303530-A2.

XX PD 24-APR-2003.

XX PF 14-OCT-2002; 2002WO-GB004647.

XX PR 12-OCT-2001; 2001GB-00024593.

XX PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX PI James B, Bacon J, March P;

XX DR WPI; 2003-393501/37.

XX DR N-PSDB; ACF39375.

XX New isolated mycobacterial peptide encoded by a gene that is induced or
 FT up-regulated under high oxygen tension, useful for diagnosing, treating
 PT or preventing a mycobacterial infection.

XX Claim 1; Page 233-234; 392pp; English.

XX The present invention describes an isolated mycobacterial peptide (I), or
 CC its fragment, variant or derivative encoded by a gene whose expression is
 CC induced or up-regulated during culture of a mycobacterium under
 CC continuous culture conditions of a dissolved oxygen tension of at least
 CC 30% air saturation measured at 37 plus degrees Celsius when compared with
 CC a dissolved oxygen tension of up to 10% air saturation measured at 37
 CC plus degrees Celsius. (I) has tuberculostatic activity and can be used in
 CC vaccines. The mycobacterial peptide (I) or its fragment, variant or
 CC derivative, inhibitor, antibody, attenuated mycobacterium, attenuated
 CC microbial carrier, DNA sequence, DNA plasmid, RNA sequence, or RNA vector

CC from the present invention can be used for manufacturing a medicament for
 CC treating or preventing a mycobacterial infection. The peptide or its
 CC fragment, variant or derivative, the antibody, or a polynucleotide probe
 CC comprising at least 8 nucleotides, where the probe binds to at least a
 CC part of the gene, is useful for manufacturing a diagnostic reagent for
 CC identifying a mycobacterial infection. The present sequence represents a
 CC Mycobacterium tuberculosis mycobacterial antigen, which is used in the
 CC exemplification of the present invention

XX Sequence 399 AA;

Query Match 6.8%; Score 8; DB 6; Length 399;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALA 25
 Db 160 GGVLAALA 167
 |||||

RESULT 659

AD117251
 ID AD117251 standard; protein; 419 AA.

XX AC AD117251;

XX DT 15-APR-2004 (first entry)

XX DE Pig NOVX protein homologue SeqID 787.

XX KW pig; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
 KW inflammation; autoimmune disorder; allergy; blood disorder;
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
 KW Alzheimer's disease; infection; str.

XX OS Sus scrofa.

XX PN WO200268649-A2.

XX PD 06-SEP-2002.

XX PF 31-JAN-2002; 2002WO-US002785.

XX PR 31-JAN-2001; 2001US-0265395P.

XX PR 31-JAN-2001; 2001US-0265412P.

XX PR 31-JAN-2001; 2001US-0265514P.

XX PR 31-JAN-2001; 2001US-0265517P.

XX PR 02-FEB-2001; 2001US-0266406P.

XX PR 05-FEB-2001; 2001US-0266767P.

XX PR 07-FEB-2001; 2001US-0266975P.

XX PR 07-FEB-2001; 2001US-0267057P.

XX PR 08-FEB-2001; 2001US-0267459P.

XX PR 09-FEB-2001; 2001US-0267823P.

XX PR 15-FEB-2001; 2001US-0268974P.

XX PR 26-FEB-2001; 2001US-0271664P.

XX PR 27-FEB-2001; 2001US-0271839P.

XX PR 27-FEB-2001; 2001US-0271855P.

XX PR 02-MAR-2001; 2001US-0272788P.

XX PR 02-MAR-2001; 2001US-0273046P.

XX PR 14-MAR-2001; 2001US-0275925P.

XX PR 14-MAR-2001; 2001US-0275947P.

XX PR 14-MAR-2001; 2001US-0275950P.

XX PR 14-MAR-2001; 2001US-0275989P.

XX PR 15-MAR-2001; 2001US-0276448P.

XX PR 16-MAR-2001; 2001US-0276450P.

XX PR 16-MAR-2001; 2001US-0276397P.

XX PR 20-MAR-2001; 2001US-0278652P.

XX PR 26-MAR-2001; 2001US-0278775P.

XX PR 26-MAR-2001; 2001US-0278778P.

XX PR 29-MAR-2001; 2001US-0279882P.

XX PR 29-MAR-2001; 2001US-0279884P.

PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 21-AUG-2001; 2001US-0312908P.
PR 28-AUG-2001; 2001US-0313390P.
PR 31-AUG-2001; 2001US-0315470P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
PA (CURA-) CURAGEN CORP.
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
XX Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RV, Gusev VT, Colman SD, Wolenc AR, Pena CEA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DW, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 787; 1498pp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytoskeletal, cardiac, antiinflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 419 AA;

Query Match 6.8%; Score 8; DB 5; Length 419;
Best Local Similarity 100.0%; Pred. No. 62;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 15 VLLGGVLA 22
|||||
Db 9 VLLGGVLA 16
RESULT 660
AAW31885
ID AAW31885 standard; protein; 465 AA.
XX
AC AAW31885;
XX
DT 03-FEB-1998 (first entry)
XX
DE HCV-1 nonstructural protein 3 (NS3) helicase fragment.
XX
KW Hepatitis C virus; HCV; nonstructural protein 3; NS3; fusion protein;
KW HCV-1 polypeptide; serine type proteinase; NTPase/RNA helicase;
KW trypsin like serine proteinase; inhibitory activity; anti-HCV agent;
KW truncated protein; colourimetric assay; metal-chelate chromatography.
XX
OS Hepatitis C virus type 1.
XX
FH Key Location/Qualifiers
FT Misc-difference 79 /note= "preferentially Pro but can be Leu"
FT Misc-difference 258 /note= "preferentially Cys but can be Tyr"
FT Misc-difference 276 /note= "preferentially Thr but can be Ser"
XX
FN WO9712043-A2.
XX
PD 03-APR-1997.
XX
PF 12-SEP-1996; 96WO-US014688.
XX
PR 15-SEP-1995; 95US-00529169.
XX
PA (CHIR) CHIRON CORP.
XX
PI Hang J, Choe J;
XX
DR WPI; 1997-212902/19.
XX
PT Soluble fragment of hepatitis C virus NS3 helicase and related fusion
PT proteins - useful for screening cpds. for anti-HCV activity.
XX
PS Claim 1; Page 8-9; 40pp; English.
XX
CC This sequence represents a hepatitis C virus (HCV) nonstructural protein
CC 3 (NS3) helicase fragment, comprising amino acids 167-631 of the NS3
CC protein AAW31884. NS3 has a conserved sequence of motifs of a serine type
CC proteinase, and of a nucleoside triphosphate (NTPase)/RNA helicase. One
CC third of the N-terminal of the HCV NS3 protein has been shown to be a
CC trypsin like serine proteinase. Two thirds of the NS3 C-terminal fragment
CC has been shown to encode NTPase activity. The helicase fragment of
CC NS3, comprising amino acids 1193-1657 of the NS3 protein, is used to
CC screen compounds for specific inhibitory activity, i.e. to identify
CC potential anti-HCV agents. The N and C termini of the present sequence
CC are putative, the actual termini being defined by expressing and
CC processing in an appropriate host. Truncated NS3 helicase fragment
CC analogues and NS3 helicase mutants have also been produced. Helicase
CC fragments with N-terminal truncations (up to 20 amino acids) are soluble
CC in purification and assay buffers, unlike full-length NS3 protein.
CC Expression of a soluble protein eliminates the need for denaturing and
CC refolding, and also provides higher yields. Expression of NS3 as a fusion
CC protein allows products to be quantified by colourimetric assay and
CC purified by metal-chelate chromatography
XX
SQ Sequence 465 AA;

Query Match 6.8%; Score 8; DB 2; Length 465;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
 Db 456 CMSADLEV 463

RESULT 661

AAM50076
 ID AAM50076 standard; protein; 473 AA.

XX AAM50076;

XX 07-AUG-2003 (revised)

DT 12-AUG-2002 (first entry)

XX HCV NS3 helicase protein.

XX Helicase NS3; protein co-ordinate date; UHCV-A; UHCV-B; UHCV-C; virucide;

KW drug discovery; inhibitor design.

XX Hepatitis C virus.

Key Location/Qualifiers

FT Domain 33..39

FT /label= beta_sheet_1

FT Region 39..48

FT /note= "motif I"

FT Domain 44..55

FT /label= alpha_helix_1

FT Domain 59..64

FT /label= beta_sheet_2

FT Region 62..73

FT /note= "motif Ia"

FT Domain 67..81

FT /label= alpha_helix_2

FT Domain 86..89

FT /label= beta_sheet_3

FT Domain 92..41

FT /label= beta_sheet_4

FT Domain 99..104

FT /label= beta_sheet_5

FT Domain 105..111

FT /label= alpha_helix_3

FT Domain 120..125

FT /label= beta_sheet_6

FT Region 125..131

FT /note= "motif II"

FT Domain 131..145

FT /label= alpha_helix_4

FT Domain 151..157

FT /label= beta_sheet_7

FT Region 157..170

FT /note= "motif III"

FT Domain 171..176

FT /label= beta_sheet_1'

FT Domain 191..195

FT /label= alpha_helix_1'

FT Domain 198..203

FT /label= beta_sheet_2'

FT Region 200..208

FT /note= "motif IV"

FT Region 205..209

FT /note= "domain 1/domain 2 interface"

FT Domain 206..219

FT /label= alpha_helix_2'

FT Domain 222..226

FT /label= beta_sheet_3'

FT Binding-site 230..232

FT /note= "domain 1 oligonucleotide binding site"

FT Region 232..238

FT Domain /note= "domain 1/domain 2 interface"
 FT 241..246
 FT /label= beta_sheet_5'
 FT 247..253
 FT /label= alpha_helix_3'
 FT 255
 FT /note= "domain 1 oligonucleotide binding site"
 FT 258..263
 FT /label= beta_sheet_6'
 FT 266..272
 FT /label= beta_sheet_6a'
 FT 269
 FT /note= "domain 1 oligonucleotide binding site"
 FT 270..271
 FT /note= "domain 1 oligonucleotide binding site"
 FT 280..286
 FT /label= beta_sheet_6b'
 FT 289..297
 FT /label= alpha_helix_4'
 FT 306..311
 FT /label= beta_sheet_7'
 FT 323..336
 FT /label= alpha_helix_5
 FT 340..353
 FT /label= alpha_helix_6
 FT 364..374
 FT /label= alpha_helix_7
 FT 378..388
 FT /label= alpha_helix_8
 FT 391..393
 FT /note= "domain 2 oligonucleotide binding site"
 FT 392..405
 FT /label= alpha_helix_9
 FT 411..413
 FT /note= "domain 2 oligonucleotide binding site"
 FT 415..427
 FT /label= alpha_helix_10
 FT 415..420
 FT /note= "domain 1/domain 2 interface"
 FT 415
 FT /note= "domain 2 oligonucleotide binding site"
 FT 416
 FT /note= "domain 2 oligonucleotide binding site"
 FT 448..462
 FT /label= alpha_helix_11
 FT 460..467
 FT /note= "domain 1/domain 2 interface"
 FT 460
 FT /note= "domain 2 oligonucleotide binding site"
 FT WO200188113-A2.
 PN 22-NOV-2001.
 PD 02-MAY-2001; 2001WO-US014233.
 PF 03-MAY-2000; 2000US-0201598P.
 PR (PRAA) PHARMACIA & UPJOHN CO.
 PA Finzel BC, Harris MS, Baldwin E;
 PI WPI; 2002-415245/44.
 DR Novel molecule/molecular complex for solving structures of other
 XX molecules/complexes, comprises portion of Hepatitis C virus helicase or
 PT helicase-like domain 1/domain 2 interface or oligonucleotide binding
 PT site.
 XX Claim 15; Fig 2; 255pp; English.
 PS This invention describes a novel molecule or complex (I) comprising a
 XX portion of Hepatitis C virus helicase or helicase-like domain 1/domain 2
 CC

CC interface (D) or oligonucleotide binding site (O) where (D) or (O) is
 CC defined by a set of points having a root mean square deviation of less
 CC than about 1.5 Angstrom from points representing the backbone atoms of
 CC the amino acids as represented by the protein structure coordinate data
 CC of UHCV-A, UHCV-B, or UHCV given in the specification. The products of
 CC the invention have virucide activity and inhibit HCV helicase activity.
 CC The method described in the invention is useful for solving the
 CC structures of other molecules or molecular complexes, in drug discovery,
 CC for designing inhibitors of HCV helicase-like binding sites, and for
 CC solving the structure of HCV helicase, HCV helicase mutants or HCV
 CC helicase homologs co-complexed with a variety of chemical entities.
 CC This sequence represents the HCV NS3 helicase described in the method of
 CC the invention. (Updated on 07-AUG-2003 to correct OS field.)

XX Sequence 473 AA;

Query Match 6.8%; Score 8; DB 5; Length 473;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
 Db 457 CMSADLEV 464
 |||||

RESULT 662

AAR29849
 ID AAR29849 standard; protein; 477 AA.

XX AAR29849;

XX 25-MAR-2003 (revised)
 DT 26-APR-1993 (first entry)

XX HCV NS2-NS4 peptide N16N15.

XX Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum; HC;
 KW transcripase; cDNA; primer; allele.

XX Hepatitis C virus.

XX Key Location/Qualifiers
 FH Misc-difference 90 /label= Pro, Ser
 FT Misc-difference 184 /label= Tyr, His
 FT Misc-difference 192 /label= Glu, Lys
 FT Misc-difference 194 /label= Thr, Ala
 FT Misc-difference 202 /label= Tyr, Phe
 FT Misc-difference 205 /label= Thr, Ala

XX EP518313-A2.

FN 16-DEC-1992.

XX 11-JUN-1992; 92EP-00109812.

XX 11-JUN-1991; 91JP-00139268.

PR 12-JUL-1991; 91JP-00172794.

PR 07-OCT-1991; 91JP-00287008.

PR 16-DEC-1991; 91JP-00332329.

PR 20-APR-1992; 92JP-00099957.

XX (MITU) MITSUBISHI KASEI CORP.

XX Seki M, Honda Y, Takahashi K, Murakami T, Teranishi Y, Hayashi N;

XX WPI; 1992-417213/51.

DR N-PSDB; AAQ32480.

XX New hepatitis C virus gene and its encoded protein - used for diagnosing
 PT and vaccinating against hepatitis C virus infections.

XX Disclosure; Page 131-34; 305pp; English.

XX The sequences given in AAR29660, AAR29559-60 and AAR29843-51 were encoded
 CC by clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV)
 CC gene of the invention. These sequences were isolated from the serum of a
 CC patient suffering from hepatitis C (HC). The NS2-NS4 RNA sequences were
 CC converted into cDNA using transcriptase in the presence of one of the
 CC primer sequences given in AAQ32553-64. The cDNA sequences were then
 CC amplified using primer pairs. The cDNA sequences isolated represent
 CC different alleles of the same region of the HCV gene. Sequence
 CC comparisons of these clones showed that it is possible for a patient to
 CC carry more than one HCV strain at one time. See also AAQ32436. (Updated
 CC on 25-MAR-2003 to correct PN field.)

XX Sequence 477 AA;

Query Match 6.8%; Score 8; DB 2; Length 477;
 Best Local Similarity 100.0%; Pred. No. 69;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9

Db 206 CMSADLEV 213
 |||||

RESULT 663

AAB69031
 ID AAB69031 standard; peptide; 512 AA.

XX AAB69031;

XX 17-APR-2001 (first entry)

XX HCV recombinant antigen pHCV-23 amino acid sequence SEQ ID NO:58.

XX Hepatitis C virus; HCV; antigen; detection; antibody.

XX Hepatitis C virus.

XX US6172189-B1.

XX 09-JAN-2001.

XX 02-JUN-1997; 97US-00867611.

XX 24-AUG-1990; 90US-00572822.

PR 07-NOV-1990; 90US-00614069.

PR 21-AUG-1991; 91US-00748561.

PR 21-AUG-1991; 91US-00748565.

PR 21-AUG-1991; 91US-00748566.

PR 19-NOV-1992; 92US-00989843.

PR 10-JAN-1994; 94US-00179896.

PR 01-MAY-1996; 96US-00646757.

XX (ABBO) ABBOTT LAB.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2001-122352/13.

XX New recombinant antigens representing distinct antigenic regions of
 PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
 PT antigens in body fluids of individuals exposed to HCV.

XX Claim 4; Col 205-208; 167pp; English.

XX The present invention describes recombinant Hepatitis C virus (HCV)
 CC antigens (I). (I) is useful as a reagent for the detection of antibodies

CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
 CC uses reliable and efficient reagents and methods to accurately detect the
 CC presence of HCV antibodies in samples obtained from individuals suspected
 CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
 CC AAB69001 to AAB69032 represent sequences used in the exemplification of
 CC the present invention

XX Sequence 512 AA;

Query Match 6.8%; Score 8; DB 4; Length 512;

Best Local Similarity 100.0%; Pred. No. 73;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67

Db 277 DEMEECSQ 284

RESULT 664

ABW01896
 ID ABW01896 standard; protein; 512 AA.

XX AC ABW01896;

XX DT 12-FEB-2004 (first entry)

XX HCV CKS-BCD recombinant antigen, pHCV-23.

XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
 KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
 KW fusion protein.

XX Chimeric - Hepatitis C virus.

OS Chimeric - Escherichia coli.

OS Chimeric - Unidentified.

XX US6593083-B1.

XX 15-JUL-2003.

XX 17-OCT-2000; 2000US-00690359.

XX 24-AUG-1990; 90US-00572822.

PR 07-NOV-1990; 90US-00614069.

PR 21-AUG-1991; 91US-00748561.

PR 21-AUG-1991; 91US-00748565.

PR 21-AUG-1991; 91US-00748566.

PR 10-NOV-1992; 92US-00989843.

PR 10-JAN-1994; 94US-00179896.

PR 01-MAY-1996; 96US-00646757.

PR 02-JUN-1997; 97US-00867611.

XX (DEVA/) DEVARE S. G.

PA (DESA/) DESAI S. M.

PA (CASE/) CASEY J. M.

PA (DAIL/) DAILEY S. H.

PA (DAWS/) DAWSON G. J.

PA (GUTI/) GUTIERREZ R. A.

PA (LESN/) LESNIEWSKI R. R.

PA (STEW/) STEWART J. L.

PA (RUPP/) RUPPRECHT K. R.

XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;

PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;

XX WPI; 2003-828264/77.

XX Identifying the presence of an antibody in a fluid sample, where the

PT antibody is immunologically reactive with a Hepatitis C virus (HCV)

PT antigen by contacting the fluid sample, which may contain a HCV antibody

PT with at a polypeptide.

XX Claim 1; Col 205-208; 168pp; English.

XX

CC The invention relates to a method for identifying the presence of an
 CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
 CC The method involves providing a fluid sample containing at least one HCV
 CC antibody, contacting the fluid sample with at least one polypeptide or
 CC recombinant fusion protein for complexing the antibody with the
 CC polypeptide or recombinant fusion protein to provide an antibody-
 CC polypeptide complex and detecting the complex. The present sequence is
 CC pHCV-23 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
 CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
 CC enzyme, linker and HCV (non- structural region) NS4 region. This sequence
 CC is used in the invention

XX Sequence 512 AA;

Query Match 6.8%; Score 8; DB 7; Length 512;

Best Local Similarity 100.0%; Pred. No. 73;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67

Db 277 DEMEECSQ 284

RESULT 665

AAE19893

ID AAE19893 standard; protein; 613 AA.

XX AC AAE19893;

XX 18-JUN-2002 (first entry)

XX Hepatitis C virus (HCV) NS3 protein.

DE Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;

XX cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer.

KW Hepatitis C virus.

XX WO200213855-A2.

XX 21-FEB-2002.

XX 15-AUG-2001; 2001WO-IB001808.

XX 17-AUG-2000; 2000US-0225767P.

PR 29-AUG-2000; 2000US-0229175P.

PR 03-NOV-2000; 2000US-00705547.

XX (TRIP-) TRIPEP AB.

PA Sallberg M, Hultgren C;

XX WPI; 2002-241837/29.

XX Vaccine compositions for treating and preventing disease, preferably

PT hepatitis C virus infection, comprises ribavirin and antigen that has

PT epitope present in hepatitis C virus.

XX Claim 11; Page 75-77; 120pp; English.

XX The invention relates to a composition comprising ribavirin and an

CC antigen preferably non structural 3 protein (NS3)/4A fragment of

CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV

CC sequence. The composition is useful for enhancing an immune response to a

CC hepatitis C antigen in humans, domestic, sport or pet species and as

CC vaccines for treating and preventing HCV infections. The composition is

CC also useful for treating viral, bacterial, fungal diseases and cancer.

CC The present sequence is HCV NS3 protein

XX Sequence 613 AA;

Query Match 6.8%; Score 8; DB 5; Length 613;

PI Hang J, Choe J;
XX WPI; 1997-212902/19.
XX
XX Soluble fragment of hepatitis C virus NS3 helicase and related fusion
PT proteins - useful for screening cpds. for anti-HCV activity.
XX
XX Disclosure; Fig 1; 40pp; English.
XX
XX This sequence represents one of the hepatitis C virus (HCV) nonstructural
CC protein 3 (NS3) molecules. It is approximately 1027-1657 amino acids of
CC the HCV-1 polyprotein. This protein has a conserved sequence of motifs of
CC a serine type proteinase, and of a nucleoside triphosphate (NTPase)/RNA
CC helicase. One third of the N-terminal of the HCV NS3 protein has been
CC shown to be a trypsin like serine proteinase. Two thirds of the NS3 C-
CC terminal fragment has been shown to encode NTPase/RNA activity. The
CC helicase fragment of NS3, comprising amino acids 1193-1657 of the NS3
CC protein, is used to screen compounds for specific inhibitory activity,
CC i.e. to identify potential anti-HCV agents. Helicase fragments with N-
CC terminal truncations (up to 20 amino acids) are soluble in purification
CC and assay buffers, unlike full-length NS3 protein. Expression of a
CC soluble protein eliminates the need for denaturing and refolding, and
CC also provides higher yields. Expression of NS3 as a fusion protein allows
CC products to be quantified by colourimetric assay and purified by metal-
CC chelate chromatography
XX
XX Sequence 631 AA;
SQ

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CWSADLEV 9
Db ||||| 622 CWSADLEV 629

RESULT 669
AAW93482
ID AAW93482 standard; protein; 631 AA.
XX
XX AAW93482;
AC
XX 11-JUN-1999 (first entry)
DE HCV NS3 protein.
XX
XX NS3; helicase domain; X-ray crystal structure; Hepatitis C virus; HCV;
KW computer programme; binding pocket; three-dimensional.
XX
XX Hepatitis C virus.
XX
XX WO9909148-A1.
PN
XX 25-FEB-1999.
XX
XX 13-AUG-1998; 98WO-US016879.
XX
XX 13-AUG-1997; 97US-0055772P.
PR
XX (VERT-) VERTEX PHARM INC.
PA
XX Kim JL, Morgenstern K, Caron P, Lin C;
PI WPI; 1999-190157/16.
XX N-PsDB; AAX23258.
DR
XX New hepatitis C virus NS3 helicase crystals - which provide molecular
PT design techniques to identify, select and design agents which bind to the
PT helicase, particularly inhibitor compounds.
XX
XX Disclosure; Page 202-205; 224pp; English.
PS
XX

CC This invention relates to the X-ray crystal structure of the Hepatitis C
CC virus helicase domain. The invention describes crystallized complexes of
CC HCV helicase and an oligonucleotide. The described method is used in a
CC novel computer programme where the computer is programmed with the
CC structure coordinates of the HCV helicase oligonucleotide binding pocket
CC or the HCV helicase nucleotide triphosphate pocket wherein the computer
CC is capable of displaying a three-dimensional representation of that
CC binding pocket
XX
XX Sequence 631 AA;
SQ

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CWSADLEV 9
Db ||||| 622 CWSADLEV 629

RESULT 670
ADV86617
ID ADV86617 standard; protein; 631 AA.
XX
XX ADV86617;
AC
XX 10-MAR-2005 (first entry)
DT
XX
XX HCV 1a strain NS3 helicase protein.
DE
XX Screening; helicase; drug discovery.
KW
XX Hepatitis C virus.
OS
XX
XX Key Location/Qualifiers
FH Region 1..180 /note= "Protease activity"
FT Domain 181..327
FT /note = Helicase subdomain I, also includes NTPase
FT binding subdomain
FT Domain 328..483
FT /note = Helicase subdomain II
FT Region 336..339
FT /note= "Beta-sheet conformation"
FT Region 347..349
FT /note= "Beta-sheet and in alpha-helical conformation"
FT Region 353..359
FT /note= "Beta-sheet conformation"
FT Region 356..359
FT /note= "Beta-sheet and in alpha-helical conformation"
FT Region 363..367
FT /note= "Beta-sheet conformation"
FT Region 371..381
FT /note= "Alpha-helical conformation"
FT Region 387..391
FT /note= "Beta-sheet conformation"
FT Region 406..411
FT /note= "Beta-sheet conformation"
FT Region 424..427
FT /note= "Beta-sheet conformation"
FT Region 431..451
FT /note= "Antiparallel beta-loop"
FT Region 455..462
FT /note= "Alpha-helical conformation"
FT Domain 460..468
FT /note = RNA binding subdomain
FT Region 471..477
FT /note= "Beta-sheet conformation"
FT Domain 484..631
FT /note = Helicase subdomain III, also includes alpha
FT helical subdomain
XX
XX US2004253577-A1.

XX PD 16-DEC-2004.
XX PF 03-APR-2001; 2001US-00825423.
XX PP 04-APR-2000; 2000US-0194419P.
XX PR (WEBER/) WEBER P C.
XX PA (REIC/) REICHERT P.
XX PA (MADI/) MADISON V S.
XX PA (WYSS/) WYSS D F.
XX PA (YAON/) YAO N.
XX PA (LIUD/) LIU D.
XX PA (GESE/) GESELL J J.
XX PI Weber PC, Reichert P, Madison VS, Wyss DF, Yao N, Liu D;
XX PI Gesell JJ;
XX PT WPI; 2005-030206/03.
XX DR
XX CC Novel polypeptide fragment of hepatitis C virus (HCV) helicase protein
XX PT derived from subdomain of HCV NS3 helicase, useful for NMR-based drug
XX PT discovery techniques, for evaluating mechanism of action and substrates
XX PT for HCV NS3 helicase.
XX PS Example 1; SEQ ID NO 1; 29pp; English.
XX CC The present invention relates to a polypeptide fragment of a hepatitis C
XX CC virus (HCV) helicase protein, which is derived from a subdomain of the
XX CC HCV NS3 helicase protein, where the fragment is less than 30 kDa,
XX CC structurally sound, soluble, monodisperse, and stable in a buffered
XX CC solution. The invention is useful for identifying an inhibitor compound
XX CC of an HCV helicase protein, for NMR-based drug discovery techniques, for
XX CC evaluating the mechanism of action and substrates for HCV NS3 helicase
XX CC and for probing NTP and nucleic acid binding sites of HCV NS3 helicase
XX CC using NMR spectroscopy and X-ray crystallography techniques. The present
XX CC sequence is the HCV 1a strain NS3 helicase protein.
XX SQ Sequence 631 AA;
Query Match 6.8%; Score 8; DB 9; Length 631;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 671
ID AAE21847 standard; protein; 632 AA.
AC AAE21847;
XX 16-JUL-2002 (first entry)
XX Hepatitis C virus NS3/4A truncated mutant protein #1.
XX Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
XX virucide; mutant; mutin.
XX Synthetic.
XX WO200214362-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001774.
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX

PR 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M;
XX WPI; 2002-339446/37.
XX Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
XX PT absence of hepatitis C virus in a subject and for preparing a medicament
XX PT for treating hepatitis C virus infection.
XX PS Example 1; Page 83-84; 90pp; English.
XX CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
XX CC proteins and their corresponding polynucleotides. NS3/4A sequences are
XX CC useful for identifying the presence or absence HCV in a subject. They are
XX CC useful for preparing a medicament used for treating or preventing HCV
XX CC infection. Sequences of the invention are also used as vaccines. The
XX CC present sequence is HCV NS3/4A truncated mutant protein
XX SQ Sequence 632 AA;
Query Match 6.8%; Score 8; DB 5; Length 632;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 623 CMSADLEV 630
RESULT 672
ID AAE19905 standard; protein; 632 AA.
XX AAE19905;
XX AAE19905;
XX 18-JUN-2002 (first entry)
XX Hepatitis C virus (HCV) NS3 protein.
XX Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX cytosolic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX Hepatitis C virus.
XX WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001808.
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX Vaccine compositions for treating and preventing disease, preferably
XX PT hepatitis C virus infection, comprises ribavirin and antigen that has
XX PT epitope present in hepatitis C virus.
XX Example 6; Page 99-101; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
XX CC antigen preferably non structural 3 protein (NS3)/4A fragment of
XX CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX CC sequence. The composition is useful for enhancing an immune response to a

CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS3 protein
XX
SQ Sequence 632 AA;

Query Match 6.8%; Score 8; DB 5; Length 632;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
DB 623 CMSADLEV 630

RESULT 673
ABW00356
ID ABW00356 standard; protein; 632 AA.
XX
AC ABW00356;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A protein fragment.
XX
KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide.
XX
OS Hepatitis C virus.
XX
PN US2002136740-A1.
XX
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PA (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
PI Sallberg M, Hultgren C;
XX
DR WPI; 2003-764978/72.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Example 6; Page 65-66; Opp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A protein fragment
XX
SQ Sequence 632 AA;

Query Match 6.8%; Score 8; DB 7; Length 632;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
DB 623 CMSADLEV 630

RESULT 674
ABW00356
ID ABW00356 standard; protein; 632 AA.
XX
AC ABW00356;
XX
DT 15-JAN-2004 (first entry)
XX
DE Hepatitis C virus NS3/4A protein fragment.
XX
KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide.
XX
OS Hepatitis C virus.
XX
PN US2002136740-A1.
XX
PD 26-SEP-2002.
XX
PF 15-AUG-2001; 2001US-00929955.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PA (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
PI Sallberg M, Hultgren C;
XX
DR WPI; 2003-764978/72.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Example 6; Page 65-66; Opp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis C
CC virus NS3/4A protein fragment
XX
SQ Sequence 632 AA;

Query Match 6.8%; Score 8; DB 7; Length 632;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
DB 623 CMSADLEV 630

RESULT 675
AAE10068
ID AAE10068 standard; protein; 638 AA.
XX
AC AAE10068;
XX
DT 29-NOV-2001 (first entry)
XX
DE Hepatitis C virus (HCV) NS3 protease la(H) protein.
XX
KW Hepatitis C virus; HCV protease; chromogenic substrate; fluorophore;
XX chromophore; fluorogenic; fluorescence polarisation substrate.
XX
OS Hepatitis C virus.
XX
PN US6251583-B1.
XX
PD 26-JUN-2001.

ADG47669
ID ADG47669 standard; protein; 632 AA.
XX
AC ADG47669;
XX
DT 11-MAR-2004 (first entry)
XX
DE HCV NS3/4A domain mutant #10.
XX
KW immunogen; hepatitis C virus; HCV infection; vaccine; mutant; mutein.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
PN US2003206919-A1.
XX
PD 06-NOV-2003.
XX
PF 26-NOV-2002; 2002US-00307047.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 15-AUG-2001; 2001US-00929955.
XX
PR 15-AUG-2001; 2001US-00930591.
XX
PA (SALL/) SALLBERG M.
XX
PI Sallberg M;
XX
DR WPI; 2004-051480/05.
XX
PT New purified or isolated nucleic acid useful for enhancing an immune
PT response to a hepatitis C antigen comprises specific nucleotide sequences
PT and the amino acid sequences.
XX
PS Example 1; SEQ ID NO 12; 83pp; English.
XX
CC The invention relates to a purified or isolated nucleic acid. The
CC peptides are useful as immunogens for the treatment and prevention of
CC hepatitis C virus (HCV) infection, in vaccine and immunogen compositions.
CC The nucleic acid and the peptide enhance an immune response to a
CC hepatitis C antigen and are potent immunogens. The present sequence is
XX used in the exemplification of the invention.
XX
SQ Sequence 632 AA;

Query Match 6.8%; Score 8; DB 8; Length 632;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
DB 623 CMSADLEV 630

RESULT 675
AAE10068
ID AAE10068 standard; protein; 638 AA.
XX
AC AAE10068;
XX
DT 29-NOV-2001 (first entry)
XX
DE Hepatitis C virus (HCV) NS3 protease la(H) protein.
XX
KW Hepatitis C virus; HCV protease; chromogenic substrate; fluorophore;
XX chromophore; fluorogenic; fluorescence polarisation substrate.
XX
OS Hepatitis C virus.
XX
PN US6251583-B1.
XX
PD 26-JUN-2001.

XX 08-APR-1999; 99US-00288391.
PF 27-APR-1998; 98US-0083204P.
PR (SCHE) SCHERING CORP.
XX Zhang R, Malcolin BA, Beyer BM, Njoroge FG, Durkin JP, Windsor WT;
PI WPI; 2001-556521/62.
XX New chromogenic, fluorogenic and fluorescence polarization hepatitis C
XX virus (HCV) substrates useful in HCV NS3 protease and inhibitor assays.
XX Example 1; Col 33-36; 21pp; English.
PS The invention relates to a chromogenic, fluorogenic and fluorescence
CC polarisation hepatitis C virus (HCV) substrate. The substrate comprises a
CC single chromophore or fluorophore linked to the C-terminus of a peptide
CC sequence, or a fluorescence polarisation HCV substrate comprising a
CC peptide sequence linked at opposite ends of the cleavage site to a
CC fluorophore and a high molecular weight binding group. The chromogenic,
CC fluorogenic and fluorescence polarisation peptide substrates provide
CC optimised specificity, better cleavage efficiency and improved
CC detectability. The chromogenic, fluorogenic and fluorescence polarisation
CC peptide substrates are useful in discovering inhibitors of HCV proteases,
CC in progress curve analysis for reversible and irreversible binding
CC inhibitors for the HCV NS3 protease. These substrates may also be used in
CC monitoring of inhibition kinetics and rapid characterisation of HCV NS3
CC protease inhibitors, and to aid in the classification of inhibitors
CC binding to either the S or S' pocket. The present sequence is HCV NS3
CC protease full length HCV 1a(H) non-covalent complex
XX
SQ Sequence 638 AA;
Query Match 6.8%; Score 8; DB 4; Length 638;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 629 CMSADLEV 636
|||||||
RESULT 676
ABG15384
ID ABG15384 standard; protein; 724 AA.
XX
AC ABG15384;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #15375.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX

DR N-PSDB; AAS79571.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 45743; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 724 AA;
Query Match 6.8%; Score 8; DB 4; Length 724;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 107 NWQKLEAF 114
Db 633 NWQKLEAF 640
|||||||
RESULT 677
ABG17531
ID ABG17531 standard; protein; 724 AA.
XX
AC ABG17531;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #17522.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS81718.
XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

PS Claim 20; SEQ ID NO 47890; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 724 AA;

Query Match 6.8%; Score 8; DB 4; Length 724;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 NWQKLEAF 114
 |||||
 Db 633 NWQKLEAF 640

RESULT 678

AA222208
 ID AAR22208 standard; protein; 781 AA.

XX AAR22208;

DT 27-AUG-2003 (revised)
 DT 23-JUL-1992 (first entry)

XX Sequence of fusion protein HCV CKS-33-BCD recombinant antigen encoded by
 DE pHCV-31.

XX CKS fusion protein; antigen.

XX Hepatitis B virus.

OS EP475182-A.

PN 18-MAR-1992.

XX 26-AUG-1991; 91EP-00114296.

FR 24-AUG-1990; 90US-00573103.

XX (ABBO) ABBOTT LAB.

PA Bolling TJ, Mandeckl W;

XX WPI; 1992-089871/12.

DR N-PSDB; AAQ22963.

XX Hepatitis C virus-CKS fusion protein - produced using vector contg. lac

PT control region.

XX Example; Fig 8; 40pp; English.

XX CKS methods of protein synthesis are known in the art, for inst. a CKS
 CC method has been disclosed in EP-A- 0331 961. Six individual nucleotides
 CC representing AAs 11-150 of the HCV genome were ligated together and
 CC cloned as a 466 base pair EcoRI-BamHI fragment into the CKS fusion vector
 CC pJ0200 to produce pHCV-34. The resultant fusion protein HCV CKS-Core,
 CC consists of 239 AAs of CKS, 7 AAs contributed by linker DNA sequences,
 CC and the first 150 AAs of HCV. To construct the plasmid pHCV-31, the 781
 CC base pair EcoRI-BamHI fragment from pHCV-23 representing the HCV-BCD
 CC region was linker-adapted and ligated into pHCV-29. The resulting
 CC plasmid, designated pHCV-31, expresses HCV CKS-33-BCD recombinant antigen
 CC which consists of 239 AAs of CKS, 8 AAs contributed by linker DNA
 CC sequences, 266 AAs of the HCV NS3 region (AAs 1192-1457), 2 AAs
 CC contributed by linker DNA sequences, 256 AAs of the HCV NS4 region (AAs
 CC 1676-1931), and 10 additional AAs contributed by linker DNA sequences.
 CC (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 781 AA;

Query Match 6.8%; Score 8; DB 2; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMBECSQ 67

Db 546 DEMBECSQ 553

RESULT 679

AA221565
 ID AAR21565 standard; protein; 781 AA.

XX AAR21565;

DT 10-MAR-2003 (revised)
 DT 09-JUN-1992 (first entry)

XX HCV CKS-33-BCD - pHCV-31.

XX Hepatitis C virus; antigen; diagnosis; inhibitor; CMP-KDO synthase; CKS;
 KW HCV CKS-33-BCD; c100-3; NANBHV.

XX Hepatitis C virus.

OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 1..239
 FT /label= CKS

FT Region 240..247

FT /label= linker

FT Peptide 248..513

FT /label= 33

FT /note= "HCV region NS-3, amino acids 1192-1457"

FT Region 514..515

FT /label= linker

FT Peptide 516..771

FT /label= BCD

FT /note= "HCV region NS-4, amino acids 1676-1931"

FT Region 772..781

FT /label= linker

XX EP472207-A.

XX 26-FEB-1992.

XX 23-AUG-1991; 91EP-00114161.

XX 24-AUG-1990; 90US-00572822.

XX 07-NOV-1990; 90US-00614069.

XX (ABBO) ABBOTT LAB.

XX Devare SG, Desai SM, Casey JM, Dawson GJ, Lesniewski RR;
PI Dailey SH, Gutierrez RA, Stewart JL;
XX WPI; 1992-066430/09.
DR N-PSDB; AAQ21678.
XX Recombinant hepatitis C virus antigens - produced as fusion proteins and
PT representing distinct antigenic regions of the HCV genome.
XX Disclosure; Fig 11, Page 37-43; 115pp; English.
XX The protein (mol.wt. 90 kD), encoded by pHCV-31, is composed of two non-
CC contiguous coding regions located in the putative non-structural regions
CC of HCV designated NS-3 and NS-4. Clone BCD represents the C-terminal 256
CC amino acids of c100-3; the N-terminal amino acids are not represented.
CC The polypeptide represents a distinct antigenic region of the HCV genome
CC and can be used for the detection of antibodies and antigens for early
CC diagnosis of HCV infection. The polypeptide can also be used to develop
CC specific inhibitors of viral replication and for therapeutic purposes.
CC (Updated on 10-MAR-2003 to add missing OS field.)
XX SQ Sequence 781 AA;

Query Match 6.8%; Score 8; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMECSQ 67
Db 546 DEMECSQ 553

RESULT 680
AAR33632
ID AAR33632 standard; protein; 781 AA.
XX AC AAR33632;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 13-JUL-1993 (first entry)
XX HCV CKS-33-BCD fusion protein encoded by pHCV31.
XX Hepatitis C virus; NANBH; non-A, non-B hepatitis; CMP-KDO synthetase;
KW non-structural protein; pHCV-31; diagnosis; CKS fusion protein;
KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; immunoassay.
XX Hepatitis C virus; Virus.
OS
XX WO9304089-A1.
FN
XX 04-MAR-1993.
PD
XX 21-AUG-1992; 92WO-US006964.
PP
XX 21-AUG-1991; 91US-00748565.
PR
XX (ABBO) ABBOTT LAB.
PA
XX Desai SM, Dailey SH, Devare SG;
PI
XX WPI; 1993-093942/11.
DR
XX N-PSDB; AAQ38266.
DR
XX New recombinant NS5 region antigens - for hepatitis C assay for detecting
PT hepatitis C virus infections.
XX
XX Example 2; Page 50-53; 164pp; English.
PS
XX The HCV CKS-33c-BCD expression vector pHCV-31 was constructed from a
CC clone expressing the HCV CKS-BCD antigen (designated pHCV-23) and a clone

CC expressing the HCV CKS-33 antigen (designated pHCV-29). The HCV BCD
CC region was excised from pHCV-23 and inserted into pHCV-29 to produce pHCV
CC -31. The HCV CKS-33-BCD antigen consists of 239 amino acids of CKS, eight
CC amino acids contributed by linker DNA sequences, 266 amino acids of the
CC HCV NS3 region (amino acids 1192-1457), 2 amino acids contributed by
CC linker DNA sequences, 256 amino acids of the HCV NS4 region (amino acids
CC 1676-1931) and 10 additional amino acids contributed by linker DNA
CC sequences. Recombinant HCV CKS-33-BCD protein was obtained from a culture
CC of E.coli transformed by the vector pHCV-31. In a Western blot, the
CC recombinant protein reacted strongly with sera from HCV patients; normal
CC human serum did not react with any component of the pHCV-31 preparations.
CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to
CC standardise OS field.)
XX SQ Sequence 781 AA;

Query Match 6.8%; Score 8; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMECSQ 67
Db 546 DEMECSQ 553

RESULT 681
AAR33594
ID AAR33594 standard; protein; 781 AA.

XX AC AAR33594;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-JUL-1993 (first entry)
XX HCV CKS-33-BCD protein encoded by vector pHCV-31.
XX Hepatitis C Virus; non-A, non-B hepatitis virus; NANBH;
KW non-structural protein; CMP-KDO synthetase; CKS fusion protein;
KW CTP: CMP-3-deoxy-manno-octulosonate cytidyl transferase; immunoassay;
KW pHCV-31.
XX Hepatitis C virus; Virus.
OS
XX WO9304088-A1.
FN
XX 04-MAR-1993.
PD
XX 21-AUG-1992; 92WO-US007188.
PP
XX 21-AUG-1991; 91US-00748561.
PR
XX (ABBO) ABBOTT LAB.
PA
XX Dailey SH, Desai SM, Devare SG;
PI
XX WPI; 1993-093941/11.
DR
XX N-PSDB; AAQ38251.
DR
XX Hepatitis C assay using recombinant NS1 region antigens - for detecting
PT antibodies and antigen in body fluids from individuals exposed to
PT hepatitis C virus.
PT
XX Example 2; Page 60-63; 175pp; English.
PS
XX The HCV CKS-33c-BCD expression vector pHCV-31 was constructed from a
CC clone expressing the HCV CKS-BCD antigen (designated pHCV-23) and a clone
CC expressing the HCV CKS-33 antigen (designated pHCV-29). The HCV BCD
CC region was excised from pHCV-23 and inserted into pHCV-29 to produce pHCV
CC -31. The HCV CKS-33-BCD antigen consists of 239 amino acids of CKS, eight
CC amino acids contributed by linker DNA sequences, 266 amino acids of the
CC HCV NS3 region (amino acids 1192-1457), 2 amino acids contributed by
CC linker DNA sequences, 256 amino acids of the HCV NS4 region (amino acids

CC 1676-1931) and 10 additional amino acids contributed by linker DNA
 CC sequences. Recombinant HCV CKS-33-BCD protein was obtained from a culture
 CC of E.coli transformed by the vector pHCV-31. In a Western blot, the
 CC recombinant protein reacted strongly with sera from HCV patients; normal
 CC human serum did not react with any component of the pHCV-31 preparations.
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 781 AA;

Query Match 6.8%; Score 8; DB 2; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
 Db 546 DEMEECSQ 553
 |||||

RESULT 682
 AAR33574
 ID AAR33574 standard; protein; 781 AA.
 AC AAR33574;
 XX

DT 25-MAR-2003 (revised)
 DT 01-JUL-1993 (first entry)
 XX

DE HCV CKS-33-BCD protein encoded by vector pHCV-31.

XX Hepatitis C virus; NS3; C100 antigen; CKS fusion protein;
 KW CMP-KDO synthetase; immunodot assay; Non-A, non-B hepatitis.
 XX

OS Hepatitis C virus.
 XX

XX WO9304087-A1.
 PN

PD 04-MAR-1993.
 XX

PF 21-AUG-1992; 92WO-US007187.
 XX

PR 21-AUG-1991; 91US-00748566.
 XX

PA (ABBO) ABBOTT LAB.
 XX

PI Desai SM, Casey JM, Rupprecht KR, Devare SG;
 XX

DR WPI; 1993-093940/11.
 DR N-PSDB; AAQ38236.
 XX

PT Hepatitis C assay using recombinant C-100 region antigens - for detecting
 PT antibodies and antigen in body fluids from individuals exposed to
 PT hepatitis C virus.
 XX

PS Example 2; Page 87-89; 206pp; English.
 XX

SS The HCV CKS-33c-BCD expression vector pHCV-31 was constructed from a
 CC clone expressing the HCV CKS-BCD antigen (designated pHCV-23) and a clone
 CC expressing the HCV CKS-33 antigen (designated pHCV-29). The HCV BCD
 CC region was excised from pHCV-23 and inserted into pHCV-29 to produce pHCV
 CC -31. The HCV CKS-33-BCD antigen consists of 239 amino acids of CKS, eight
 CC amino acids contributed by linker DNA sequences, 266 amino acids of the
 CC HCV NS3 region (amino acids 1192-1457), 2 amino acids contributed by
 CC linker DNA sequences, 256 amino acids of the HCV NS4 region (amino acids
 CC 1676-1931) and 10 additional amino acids contributed by linker DNA
 CC sequences. (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 781 AA;

Query Match 6.8%; Score 8; DB 2; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
 Db 546 DEMEECSQ 553
 |||||

RESULT 683
 AAR52690
 ID AAR52690 standard; protein; 781 AA.
 XX AAR52690;
 AC

DT 25-MAR-2003 (revised)
 DT 10-JAN-1995 (first entry)
 XX

DE HCV CKS-33-BCD antigen.

XX Promoter; expression; lac; hepatitis; virus; protein; core; HCV; CKS;
 KW CTP-CMP-3-deoxy-D-manno-octulosonate cytidyl transferase;
 KW CMP-KDO synthetase; HIV; human immunodeficiency virus; glycoprotein;
 KW envelope protein; gp41; transmembrane protein.
 XX

OS Hepatitis virus.
 XX

PN US5312737-A.
 XX

PD 17-MAY-1994.
 XX

PF 14-FEB-1992; 92US-00835878.
 XX

PR 11-MAR-1988; 88US-00167067.
 PR 23-NOV-1988; 88US-00276263.
 PR 24-AUG-1990; 90US-00573103.
 XX

PA (ABBO) ABBOTT LAB.
 XX

PI Bolling TJ, Mandecki W;
 XX

DR WPI; 1994-159120/19.
 DR N-PSDB; AAQ62663.
 XX

PT Prodn. of recombinant hepatitis virus core protein - as fusion protein
 PT with E. coli enzyme CKS.
 XX

PS Example 12; Fig 24; 63pp; English.
 XX

CC The Hepatitis core virus protein (HCV) coding sequence was used in the
 CC construction of an expression vector designated pHCV-31. This vector
 CC allows the fusion of recombinant proteins to a CKS (CTP-CMP-3-deoxy-D-
 CC manno-octulosonate cytidyl transferase or CMP-KDO synthetase) protein.
 CC The use of the CKS protein coding sequence in the construct facilitates
 CC the detection and purification of heterologous proteins and gives fusion
 CC proteins which are expressed at high levels. The HCV CKS-33-BCD antigen
 CC expressed comprises 239 amino acids of CKS, eight amino acids contributed
 CC by linker DNA sequences, 266 amino acids of the HCV NS3 region, 2 amino
 CC acids contributed by linker DNA sequences, 256 amino acids of the HCV NS4
 CC region and 10 additional amino acids contributed by linker (Updated on 25
 CC -MAR-2003 to correct PF field.)
 XX

SQ Sequence 781 AA;

Query Match 6.8%; Score 8; DB 2; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
 Db 546 DEMEECSQ 553
 |||||

RESULT 684
 AAB51372
 ID AAB51372 standard; protein; 781 AA.
 XX

AC AAB51372;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE HCV recombinant antigen pHCV-31 protein sequence SEQ ID NO:4.
 XX
 KW Hepatitis C virus; HCV; antigen; detection; antibody.
 XX
 OS Hepatitis C virus.
 XX
 XX US6172189-B1.
 XX
 XX 09-JAN-2001.
 XX
 XX 02-JUN-1997; 97US-00867611.
 XX
 XX 24-AUG-1990; 90US-00572822.
 PR 07-NOV-1990; 90US-00614069.
 PR 21-AUG-1991; 91US-00748561.
 PR 21-AUG-1991; 91US-00748565.
 PR 21-AUG-1991; 91US-00748566.
 PR 19-NOV-1992; 92US-00989843.
 PR 10-JAN-1994; 94US-00179896.
 PR 01-MAY-1996; 96US-00646757.
 XX
 XX (ABBO) ABBOTT LAB.
 PA
 XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 PI
 XX WPI; 2001-122352/13.
 DR N-PSDB; AAF32219.
 DR
 XX
 PT New recombinant antigens representing distinct antigenic regions of
 PT Hepatitis C virus (HCV) genome, useful for detection of antibodies and
 PT antigens in body fluids of individuals exposed to HCV.
 XX
 XX Claim 3; Col 53-56; 167pp; English.
 XX
 XX The present invention describes recombinant Hepatitis C virus (HCV)
 CC antigens (I). (I) is useful as a reagent for the detection of antibodies
 CC and antigen in body fluids from individuals exposed to HCV. The HCV assay
 CC uses reliable and efficient reagents and methods to accurately detect the
 CC presence of HCV antibodies in samples obtained from individuals suspected
 CC of having HCV infection. AAF32218 to AAF32235, AAB51371 to AAB51379 and
 CC AAB69001 to AAB69032 represent sequences used in the exemplification of
 CC the present invention
 XX
 XX Sequence 781 AA;
 SQ
 Query Match 6.8%; Score 8; DB 4; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 60 DEMEECSQ 67
 Db 546 DEMEECSQ 553
 RESULT 685
 AABW01858
 ID AABW01858 standard; protein; 781 AA.
 XX
 AC AABW01858;
 XX
 XX 12-FEB-2004 (first entry)
 DT
 XX HCV-CXS-33-BCD recombinant antigen, pHCV-31.
 DE
 XX Hepatitis C virus; HCV; immunological; CMP-KDO synthase; CKS;
 KW CTP; CMP-3-deoxy-manno-octulosonate cytidyl transferase; antigen;
 KW fusion protein.
 XX

OS Chimeric - Hepatitis C virus.
 OS Chimeric - Escherichia coli.
 XX Chimeric - Unidentified.
 PN US6593083-B1.
 XX
 XX 15-JUL-2003.
 PD
 XX 17-OCT-2000; 2000US-00690359.
 PF
 XX 24-AUG-1990; 90US-00572822.
 PR 07-NOV-1990; 90US-00614069.
 PR 21-AUG-1991; 91US-00748561.
 PR 21-AUG-1991; 91US-00748565.
 PR 21-AUG-1991; 91US-00748566.
 PR 19-NOV-1992; 92US-00989843.
 PR 10-JAN-1994; 94US-00179896.
 PR 01-MAY-1996; 96US-00646757.
 PR 02-JUN-1997; 97US-00867611.
 XX
 XX (DEVA/) DEVARE S G.
 PA (DESA/) DESAI S M.
 PA (CASE/) CASEY J M.
 PA (DAIL/) DAILEY S H.
 PA (DAMS/) DAMSON G J.
 PA (GUTI/) GUTIERREZ R A.
 PA (LESN/) LESNIEWSKI R R.
 PA (STEW/) STEWART J L.
 PA (RUPP/) RUPPRECHT K R.
 XX
 XX Devare SG, Desai SM, Casey JM, Dailey SH, Dawson GJ;
 PI Gutierrez RA, Lesniewski RR, Stewart JL, Rupprecht KR;
 PI
 XX WPI; 2003-828264/77.
 DR N-PSDB; AAD63431.
 DR
 XX
 PT Identifying the presence of an antibody in a fluid sample, where the
 PT antibody is immunologically reactive with a Hepatitis C virus (HCV)
 PT antigen by contacting the fluid sample, which may contain a HCV antibody
 PT with at a polypeptide.
 XX
 XX Example 2; Col 53-56; 168pp; English.
 PS
 XX The invention relates to a method for identifying the presence of an
 CC antibody immunologically reactive with a Hepatitis C virus (HCV) antigen.
 CC The method involves providing a fluid sample containing at least one HCV
 CC antibody, contacting the fluid sample with at least one polypeptide or
 CC recombinant fusion protein for complexing the antibody with the
 CC polypeptide or recombinant fusion protein to provide an antibody-
 CC polypeptide complex and detecting the complex. The present sequence is
 CC pHCV-31 fusion protein which comprises Escherichia coli CKS (CTP: CMP-3-
 CC deoxy-manno-octulosonate cytidyl transferase or CMP-KDO synthase)
 CC enzyme, linker and HCV (non- structural region) NS3 and NS4 region. This
 CC sequence is used in the invention
 XX
 XX Sequence 781 AA;
 SQ
 Query Match 6.8%; Score 8; DB 7; Length 781;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 60 DEMEECSQ 67
 Db 546 DEMEECSQ 553
 RESULT 686
 AAP90158
 ID AAP90158 standard; protein; 1786 AA.
 XX
 AC AAP90158;
 XX
 XX 25-MAR-2003 (revised)
 DT

CC represents an HCV polymerase protein used in the scope of the invention.

XX Sequence 3010 AA;
SQ
Query Match 6.8%; Score 8; DB 9; Length 3010;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 1648 CMSADLEV 1655
RESULT 689
ADX40826
ID ADX40826 standard; protein; 3013 AA.
XX
AC ADX40826;
XX
DT 21-APR-2005 (first entry)
XX
DE HCV polymerase protein #49.
XX
KW Immune stimulation; polymerase; enzyme.
XX
OS Hepatitis C virus.
XX
PN WO2005012502-A2.
XX
PD 10-FEB-2005.
XX
PF 29-MAR-2004; 2004WO-US009510.
XX
PR 28-MAR-2003; 2003US-0458026P.
XX
PA (EPIM-) EPIMUNE INC.
XX
PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX
DR WPI; 2005-132661/14.
XX
PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
PT response comprises identifying variants of a peptide epitope 8-11 amino
PT acids in length comprising primary anchor residues of the same HLA class
PT I binding motif.
XX
PS Disclosure; Page 388-440; 458pp; English.
XX
CC The invention relates to a method of identifying a candidate peptide
CC epitope which induces an HLA class I CTL response against variants of the
CC peptide epitope, comprising identifying, from a particular antigen of an
CC infectious agent, variants of a peptide epitope comprising primary anchor
CC residues of the same HLA class I binding motif. The method is useful for
CC identifying a candidate peptide epitope, which induces an HLA class I CTL
CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX
SQ Sequence 3013 AA;
Query Match 6.8%; Score 8; DB 9; Length 3013;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 22 AALAAAYCL 29
Db 1666 AALAAAYCL 1673
RESULT 690
AAW71274
ID AAW71274 standard; peptide; 7 AA.
XX
AC AAW71274;

XX 17-NOV-1998 (first entry)
XX
DE Cleavable substrate sequence derived from NS3/4B junction.
XX
KW HCV NS3 protease; NS4A cofactor protein; HCV therapy; fusion protein.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
PN WO9837180-A2.
XX
PD 27-AUG-1998.
XX
PF 20-FEB-1998; 98WO-US003367.
XX
PR 22-FEB-1997; 97US-00804266.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Chen C, Molla A, Tripathi RL;
XX
DR WPI; 1998-467551/40.
XX
PT New hepatitis C virus fusion proteins - comprises NS3 protease and NS4A
PT co-factor, used in assays for screening for compounds for use in HCV
PT therapy.
XX
PS Claim 17; Page 6; 31pp; English.
XX
CC AAW71273-83 represent cleavable substrate peptide sequences. The
CC specification describes a fusion protein derived from the Hepatitis C
CC virus (HCV) NS3 protease and NS4A cofactor proteins (NS3/4A). A non-
CC autocleavable fusion protein of HCV NS3 protease and HCV NS4A cofactor
CC protein is produced upon expression, which is biologically active. The
CC products can be used to obtain drugs which can inhibit NS3 protease
CC activity for use in HCV therapy. They can also be used to design
CC compounds which interact with or inhibit the NS3/4A fusion proteins
XX
SQ Sequence 7 AA;
Query Match 5.9%; Score 7; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECS 66
Db 1 DEMEECS 7
RESULT 691
AAO23378
ID AAO23378 standard; peptide; 7 AA.
XX
AC AAO23378;
XX
DT 06-NOV-2003 (first entry)
XX
DE HCV NS4A peptide 13 (SeqID 30) used to identify binding to NS3 protease.
XX
KW Hepatitis C virus; HCV; serine protease inhibitor; sequelae; cirrhosis;
KW chronic active hepatitis; hepatocellular carcinoma; liver fibrosis; NS3;
KW necrosis; inflammation; bile duct change; cytostatic; antiinflammatory;
KW bovine viral diarrhoea; hepatotropic; antiviral; virucide; dengue fever;
KW NS4A.
XX
OS Hepatitis C virus.
XX
PN WO2003051910-A2.
XX
PD 26-JUN-2003.
XX
PF 13-DEC-2002; 2002WO-CA001929.

XX 14-DEC-2001; 2001US-0340574P.
 XX (JOYC/) JOYCE M.
 XX (WILL/) WILLIAMS M.
 XX (HIND/) HINDSGAUL O.
 XX (TYRR/) TYRREL D L.
 XX Joyce M, Williams M, Hindsgaul O, Tyrrel DL;
 XX WPI; 2003-607859/57.
 XX New peptides useful, e.g. in the treatment of or reduction of viral load
 XX of hepatitis C virus and associated conditions, e.g. liver fibrosis,
 XX necrosis, inflammation or bile duct changes.
 XX Example 3; Fig 7; 30pp; English.
 XX This invention relates to novel hepatitis C virus (HCV) protease
 XX inhibitors. Specifically, these inhibitors are small, hydrophobic
 XX peptides that work by affecting the activity of the HCV serine protease
 XX NS3, or preventing NS3 activation by inhibition of its co-factor NS4A.
 XX Chronic infection with HCV can lead to serious sequelae including
 XX chronic active hepatitis, cirrhosis and hepatocellular carcinoma, as well
 XX as HCV associated conditions including liver fibrosis, necrosis,
 XX inflammation or bile duct changes. The present invention describes these
 XX peptide inhibitors as virucides, and as such they can be used to inhibit
 XX HCV replication and reduce the viral load. They also have hepatotropic
 XX and antiinflammatory activity and can be described as cytostatic.
 XX Furthermore, the antiviral peptides derived from the relevant conserved
 XX NS3 or NS4A domains can be used to treat other viruses including the
 XX dengue fever virus and the bovine viral diarrhoea virus. This peptide
 XX sequence, peptide 13 (SeqID 30), is part of the C-terminal deletion
 XX library that was used to identify the minimal domain of NS4A that is
 XX required for binding to the NS3 protease, and hence identify inhibitor
 XX peptides of the invention
 XX Sequence 7 AA;
 Query Match 5.9%; Score 7; DB 6; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 32 GCWIVG 38
 DB 1 GCWIVG 7
 RESULT 692
 AAJ01353
 ID AAJ01353 standard; peptide; 8 AA.
 XX AAJ01353;
 AC AAJ01353;
 XX 02-JUL-2001 (first entry)
 XX Hepatitis C virus epitope #1344.
 XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 XX antiviral.
 XX Hepatitis C virus.
 XX WO200121189-A1.
 XX 29-MAR-2001.
 XX 19-JUL-2000; 2000WO-US019774.
 XX 19-JUL-1999; 98US-00357737.
 XX (EPIM-) EPIMUNE INC.
 XX

PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX WPI; 2001-308046/32.
 XX A new composition useful as a vaccines against hepatitis C virus.
 XX Disclosure; Page 135; 214pp; English.
 XX The present invention describes a composition comprising a prepared
 XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 XX These are derived from HCV HLA-binding motifs. They are useful in
 XX CC vaccines for the prevention and treatment of HCV infection in humans. The
 XX present sequence is an epitope used in the disclosure of the invention
 XX Sequence 8 AA;
 Query Match 5.9%; Score 7; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 MSADLEV 9
 DB 1 MSADLEV 7
 RESULT 693
 AAO23379
 ID AAO23379 standard; peptide; 8 AA.
 XX AAO23379;
 AC AAO23379;
 XX 06-NOV-2003 (first entry)
 XX HCV NS4A peptide 14 (SeqID 32) used to identify binding to NS3 protease.
 XX Hepatitis C virus; HCV; serine protease inhibitor; sequelae; cirrhosis;
 XX chronic active hepatitis; hepatocellular carcinoma; liver fibrosis; NS3;
 XX necrosis; inflammation; bile duct change; cytostatic; antiinflammatory;
 XX bovine viral diarrhoea; hepatotropic; antiviral; virucide; dengue fever;
 XX NS4A.
 XX Hepatitis C virus.
 XX WO2003051910-A2.
 XX 26-JUN-2003.
 XX 13-DEC-2002; 2002WO-CA001929.
 XX 14-DEC-2001; 2001US-0340574P.
 XX (JOYC/) JOYCE M.
 XX (WILL/) WILLIAMS M.
 XX (HIND/) HINDSGAUL O.
 XX (TYRR/) TYRREL D L.
 XX Joyce M, Williams M, Hindsgaul O, Tyrrel DL;
 XX WPI; 2003-607859/57.
 XX New peptides useful, e.g. in the treatment of or reduction of viral load
 XX of hepatitis C virus and associated conditions, e.g. liver fibrosis,
 XX necrosis, inflammation or bile duct changes.
 XX Example 3; Fig 7; 30pp; English.
 XX This invention relates to novel hepatitis C virus (HCV) protease
 XX inhibitors. Specifically, these inhibitors are small, hydrophobic
 XX peptides that work by affecting the activity of the HCV serine protease
 XX NS3, or preventing NS3 activation by inhibition of its co-factor NS4A.
 XX Chronic infection with HCV can lead to serious sequelae including
 XX chronic active hepatitis, cirrhosis and hepatocellular carcinoma, as well
 XX as HCV associated conditions including liver fibrosis, necrosis,
 XX inflammation or bile duct changes. The present invention describes these
 XX peptide inhibitors as virucides, and as such they can be used to inhibit
 XX HCV replication and reduce the viral load. They also have hepatotropic
 XX and antiinflammatory activity and can be described as cytostatic.
 XX Furthermore, the antiviral peptides derived from the relevant conserved
 XX NS3 or NS4A domains can be used to treat other viruses including the
 XX dengue fever virus and the bovine viral diarrhoea virus. This peptide
 XX sequence, peptide 13 (SeqID 30), is part of the C-terminal deletion
 XX library that was used to identify the minimal domain of NS4A that is
 XX required for binding to the NS3 protease, and hence identify inhibitor
 XX peptides of the invention
 XX Sequence 7 AA;
 Query Match 5.9%; Score 7; DB 6; Length 7;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 32 GCWIVG 38
 DB 1 GCWIVG 7
 RESULT 692
 AAJ01353
 ID AAJ01353 standard; peptide; 8 AA.
 XX AAJ01353;
 AC AAJ01353;
 XX 02-JUL-2001 (first entry)
 XX Hepatitis C virus epitope #1344.
 XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 XX antiviral.
 XX Hepatitis C virus.
 XX WO200121189-A1.
 XX 29-MAR-2001.
 XX 19-JUL-2000; 2000WO-US019774.
 XX 19-JUL-1999; 98US-00357737.
 XX (EPIM-) EPIMUNE INC.
 XX

CC as HCV associated conditions including liver fibrosis, necrosis,
 CC inflammation or bile duct changes. The present invention describes these
 CC peptide inhibitors as virucides, and as such they can be used to inhibit
 CC HCV replication and reduce the viral load. They also have hepatotropic
 CC and antiinflammatory activity and can be described as cytostatic.
 CC Furthermore, the antiviral peptides derived from the relevant conserved
 CC NS3 or NS4A domains can be used to treat other viruses including the
 CC dengue fever virus and the bovine viral diarrhoea virus. This peptide
 CC sequence, peptide 14 (SeqID 32), is part of the C-terminal deletion
 CC library that was used to identify the minimal domain of NS4A that is
 CC required for binding to the NS3 protease, and hence identify inhibitor
 CC peptides of the invention

XX SQ Sequence 8 AA;

Query Match 5.9%; Score 7; DB 6; Length 8;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
 Db 1 GCVVIVG 7
 |||||

RESULT 694

AAW82070
 ID AAW82070 standard; peptide; 9 AA.

XX AC AAW82070;

DT 18-FEB-1999 (first entry)

XX Fluorogenic protease indicator protease binding peptide #48.

DE Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
 KW conformation change.

OS Synthetic.

XX WO9837226-A1.

XX 27-AUG-1998.

XX 20-FEB-1998; 98WO-US003000.

XX 20-FEB-1997; 97US-00802981.

XX (ONCO-) ONCOIMMUNIN INC.

XX Komoriya A, Packard BS;

XX WPI; 1998-467579/40.

XX New fluorogenic compositions - containing 2 fluorophores separated by a
 PT peptide comprising a protease binding site, used for detecting protease
 PT activity in samples.

PS Claim 4; Page 77; 90pp; English.

XX AAW82023-W82240 are peptides used in the construction of a fluorogenic
 CC composition which is used for the detection of protease activity in
 CC biological samples. The products can be used for the detection of
 CC conformation changes in nucleic acids, oligosaccharides, polysaccharides,
 CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,
 CC steroids or polymers. In addition, attachment of a hydrophobic group to a
 CC molecule can be used to enhance uptake by cells. The composition is
 CC composed of P = peptide comprising a protease binding site for the
 CC protease, F1, F2 peptides = fluorophores where F1 is attached to the
 CC amino terminal amino acid and F2 is attached to the carboxyl terminal
 CC amino acid and S1, S2 peptides = when present, are peptide spacers where
 CC S1, when present, is attached to the amino terminal acid, and S2, when
 CC present, is attached to the carboxyl terminal amino acid

SQ Sequence 9 AA;

Query Match 5.9%; Score 7; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMEECSQ 67
 Db 1 EMEECSQ 7
 |||||

RESULT 695

AAJ01788
 ID AAJ01788 standard; peptide; 9 AA.

XX AC AAJ01788;

DT 02-JUL-2001 (first entry)

XX Hepatitis C virus epitope #1779.

DE Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
 KW antiviral.

XX Hepatitis C virus.

XX WO200121189-A1.

XX 29-MAR-2001.

XX 19-JUL-2000; 2000WO-US019774.

XX 19-JUL-1999; 99US-00357737.

XX (BFIM-) EPIIMUNE INC.

XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;

XX Baker DM, Cellis E, Kubo RT, Grey HM;

XX WPI; 2001-308046/32.

XX A new composition useful as a vaccines against hepatitis C virus.

XX Disclosure; Page 145; 214pp; English.

XX The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention

XX SQ Sequence 9 AA;

Query Match 5.9%; Score 7; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 Db 3 GGVLAAL 9
 |||||

RESULT 696

AAJ03112
 ID AAJ03112 standard; peptide; 9 AA.

XX AC AAJ03112;

DT 02-JUL-2001 (first entry)

XX Hepatitis C virus epitope #3103.

XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;


```
KW antiviral.
XX Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX 19-JUL-2000; 2000WO-US019774.
XX 19-JUL-1999; 99US-00357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 175; 214pp; English.
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
XX Query Match 5.9%; Score 7; DB 4; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAAL 24
XX Db 3 GGVLAAL 9
XX RESULT 698
XX ID AAJ00436 standard; peptide; 9 AA.
XX AC AAJ00436;
XX XX 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #427.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 111; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
XX Query Match 5.9%; Score 7; DB 4; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAAL 24
XX Db 3 GGVLAAL 9
XX RESULT 699
XX ID AAJ03686 standard; peptide; 9 AA.
XX XX
```

```
KW antiviral.
XX Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX 19-JUL-2000; 2000WO-US019774.
XX 19-JUL-1999; 99US-00357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 175; 214pp; English.
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX Sequence 9 AA;
XX Query Match 5.9%; Score 7; DB 4; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLAAL 24
XX Db 3 GGVLAAL 9
XX RESULT 697
XX ID AAJ00948 standard; peptide; 9 AA.
XX AC AAJ00948;
XX XX 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #939.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 124; 214pp; English.
```


fluorene-carboxylic group, 9-fluorene-carboxylic group, and 9-fluorenone-1-carboxylic group, benzoyloxy-carbonyl, Xanthyl (Xan), Trityl (Trt), 4-methyltrityl (Mtt), 4-methoxytrityl (Mmt), 4-methoxy-2,3,6-trimethylbenzenesulphonyl (Mtr), Mesitylene-2-sulphonyl (Mts), 4,4'-dimethoxybenzhydryl (Mbh), etc. The method described in the invention is useful for detecting protease or nuclease activity (or the presence of nuclease acid) in histological section, cells in culture, (e.g., seeded or cultured adherent cells), a biological sample such as tissue, biopsy, lymph, embryo, or whole animal, or cell suspension derived from a biological sample such as tissue, blood, urine, saliva, lymph, or biopsy. The indicator composition is also useful for screening a test agent for the ability to modulate a protease (or a nuclease, lipase, etc.). The indicator reagents allow rapid determination of protease activity in a matter of minutes in a single-step procedure. The fluorescent indicators both absorb and emit in the visible range (400-800 nm). These signals are therefore not readily quenched by, nor is activation of the fluorophores, that is, absorption of light, interfered with by background molecules; therefore they are easily detected in biological samples. The fluorogenic protease indicators utilise high efficiency fluorophores and are able to achieve a high degree of quenching while providing a strong signal when the quench is released by cleavage of the peptide substrate. The high signal allows detection of very low levels of protease activity. Thus the fluorogenic protease indicators are particularly well suited for in situ detection of protease activity. ABU60357-ABU60477 represent peptides use to illustrate the method described in the disclosure of the invention

SQ Sequence 9 AA;

Query Match 5.9%; Score 7; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMERCSQ 67
| | | | |
Db 1 EMERCSQ 7

RESULT 702

AAO23380
ID AAO23380 standard; peptide; 9 AA.

XX AC AAO23380;

DT 06-NOV-2003 (first entry)

DE HCV NS4A peptide 15 (SeqID 34) used to identify binding to NS3 protease.

XX HCV NS4A peptide 15 (SeqID 34) used to identify binding to NS3 protease.
XX Hepatitis C virus; HCV; serine protease inhibitor; sequellae; cirrhosis;
KW chronic active hepatitis; hepatocellular carcinoma; liver fibrosis; NS3;
KW necrosis; inflammation; bile duct change; cytostatic; antiinflammatory;
KW bovine viral diarrhoea; hepatotropic; antiviral; virucide; dengue fever;
KW NS4A.

XX OS Hepatitis C virus.

XX FN WO2003051910-A2.

XX XX 26-JUN-2003.

XX PF 13-DEC-2002; 2002WO-CA001929.

XX PR 14-DEC-2001; 2001US-0340574P.

XX PA (JOYC/) JOYCE M.

XX PA (WILL/) WILLIAMS M.

XX PA (HIND/) HINDSGAUL O.

XX PA (TYRR/) TYRREL D L.

XX PI Joyce M, Williams M, Hindsgaul O, Tyrrel DL;

XX DR WPI; 2003-607859/57.

XX PT New peptides useful, e.g. in the treatment of or reduction of viral load

PT of hepatitis C virus and associated conditions, e.g. liver fibrosis,
XX necrosis, inflammation or bile duct changes.

XX Example 3; Fig 7; 30pp; English.

XX This invention relates to novel hepatitis C virus (HCV) protease
CC inhibitors. Specifically, these inhibitors are small, hydrophobic
CC peptides that work by affecting the activity of the HCV serine protease
CC NS3, or preventing NS3 activation by inhibition of its co-factor NS4A.
CC Chronic infection with HCV can lead to serious sequelae including
CC chronic active hepatitis, cirrhosis and hepatocellular carcinoma, as well
CC as HCV associated conditions including liver fibrosis, necrosis,
CC inflammation or bile duct changes. The present invention describes these
CC peptide inhibitors as virucides, and as such they can be used to inhibit
CC HCV replication and reduce the viral load. They also have hepatotropic
CC and antiinflammatory activity and can be described as cytostatic.
CC Furthermore, the antiviral peptides derived from the relevant conserved
CC NS3 or NS4A domains can be used to treat other viruses including the
CC dengue fever virus and the bovine viral diarrhoea virus. This peptide
CC sequence, peptide 15 (SeqID 34), is part of the C-terminal deletion
CC library that was used to identify the minimal domain of NS4A that is
CC required for binding to the NS3 protease, and hence identify inhibitor
CC peptides of the invention

XX Sequence 9 AA;

Query Match 5.9%; Score 7; DB 6; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
| | | | |
Db 1 GCVVIVG 7

RESULT 703

ADZ86594
ID ADZ86594 standard; peptide; 9 AA.

XX AC ADZ86594;

DT 14-JUL-2005 (first entry)

DE Cytotoxic Hepatitis C virus T-cell epitope, SEQ ID 39.

XX Cytotoxic T-lymphocyte; antiinflammatory; virucide; hepatotropic;
KW vaccine; hepatitis C virus infection.

XX OS Hepatitis C virus.

XX FN WO2005042698-A2.

XX PD 12-MAY-2005.

XX PF 15-OCT-2004; 2004WO-US033942.

XX PR 23-OCT-2003; 2003US-0513216P.

XX PA (PECO-) PECOS LABS INC.

XX PI Lund O, Lundegaard C, Nielsen M, Worning P, Deans RJ, Buus S;
XX Brunak S;

XX DR WPI; 2005-333700/34.

XX PT New T-cell epitope, used as diagnostic tools and as vaccines/composition
XX for the treatment and prevention of hepatitis C.

XX PS Claim 1; SEQ ID NO 39; 122pp; English.

XX The invention relates to a novel cytotoxic Hepatitis C virus (HCV) T-cell
CC epitope, comprising any of the 429 fully defined amino acid sequences, or
CC their variations, given in the specification. The invention further

CC comprises: a method for predicting peptides that can be used as epitopes
 CC or as diagnostic tools; a prediction of the neural network combined with
 CC a prediction or measurement of one of the following: proteasomal cleavage
 CC sites, MHC binding, presence of sequence or related sequence(s) in patent
 CC databases, TAP binding, gene or protein expression level, function of the
 CC protein, or similarity of self proteins; and a vaccine and a diagnostic
 CC tool, each using a limited number such as at least 1-5, 8, 16, 32, 64,
 CC 128, 256, or 512 of any of the peptides stated above or any of the
 CC peptides predicted using the method above. The cytotoxic HCV 1-cell
 CC epitopes have antiinflammatory, virucide, and hepatotropic activities.
 CC The epitopes are useful as vaccines and as diagnostic tools. The vaccines
 CC and compositions are useful for treating and preventing hepatitis C virus
 CC infection. This sequence represents one of the 429 cytotoxic hepatitis C
 CC virus 1-cell epitopes of the invention.

XX
 SQ Sequence 9 AA;

Query Match 5.9%; Score 7; DB 9; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e-06;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 ALAAAYCL 29
 Db 1 ALAAAYCL 7
 |||||

RESULT 704
 AAR61536
 ID AAR61536 standard; peptide; 10 AA.

XX AAR61536;

XX 25-MAR-2003 (revised)
 DT 11-MAY-1995 (first entry)

XX Peptide fragment (1.0492) of HCV binds HLA-A2.1.

XX antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
 KW plasma specific antigen; hepatitis B virus; Epstein Barr;
 KW human immunodeficiency virus; human papilloma virus; p53; c-ERB2; MAGE-1;
 KW melanoma antigen-1; core antigen; surface antigen;
 KW pharmaceutical composition; in vivo; ex vivo; therapeutic; diagnostic;
 KW MHC class I molecule; major histocompatibility complex; HLA-A2.1; 9mer;
 KW 10mer; anchor; human leukocyte antigen.

XX Hepatitis C virus.

XX WO9420127-A1.

XX 15-SEP-1994.

XX 04-MAR-1994; 94WO-US002353.

XX 05-MAR-1993; 93US-00027146.

PR 04-JUN-1993; 93US-00073205.

PR 29-NOV-1993; 93US-00159184.

XX (CVTE-) CVTEL CORP.

XX Grey HM, Sette A, Sidney J, Kast W;

XX WPI; 1994-302678/37.

XX Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for

PT treatment or prophylaxis of cancer, virus infection or autoimmune

PT diseases.

XX Example 5; Page 108; 138pp; English.

XX AAR59496-R61666 are immunogenic 10mer peptides that contain a HLA-A2.1

CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity

CC of at least 1% as compared to a reference peptide (AAR71293). AAR61536

CC has an IC50 of 0.0084 and the sequence occurs at position 1661 in the HCV

CC LORF protein. The peptides of the invention can induce cytotoxic T
 CC lymphocytes which can react with target cells. They can be used for the
 CC treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 10 AA;

Query Match 5.9%; Score 7; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 Db 4 GGVLAAL 10
 |||||

RESULT 705

AAW13785

ID AAW13785 standard; peptide; 10 AA.

XX AAW13785;

XX 10-NOV-1997 (first entry)

DE Hepatitis C virus NS4A/NS4B junction region derivative peptide.

XX NS3; protease; HCV; inhibitor; antiviral; virucide.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label= OTHER

FT /note= "Ac-Asp"

FT Misc-difference 7 /note= "site of Ser substn."

XX WO9708304-A2.

XX 06-MAR-1997.

XX 20-AUG-1996; 96WO-IT000163.

XX 22-AUG-1995; 95IT-RM000573.

XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

XX Steinkuehler C, Peasi A, Bianchi E, Taliani M, Tomei L, Urbani A;
 XX De Francesco R, Narjes F;

XX WPI; 1997-179261/16.

XX New isolated hepatitis C virus NS3 polypeptide(s) - useful in high-
 PT throughput assays for detecting inhibitors which can be used to inhibit
 PT virus replication.

XX Claim 5; Page 46; 70pp; English.

XX A synthetic peptide (AAW13785) corresponds to a serine-substd., C- and N-
 CC terminally deleted version of the hepatitis C virus (HCV) polypeptide
 CC NS4A/4B junction (see also AAW13778). Peptides (AAW13771-86) and
 CC deipeptides (AAW13787-91) based on this junction region can be used as
 CC substrates in high throughput assays of NS3 protease activity (see also
 CC AAW13766-70). The peptides are cleaved into detectable products. The
 CC assays can be used to identify NS3 protease inhibitors and therefore
 CC therapeutic agents for use against HCV

XX Sequence 10 AA;

Query Match 5.9%; Score 7; DB 2; Length 10;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      60 DEMECS 66
DB      1 DEMECS 7

RESULT 706
AAJ01927
ID AAJ01927 standard; peptide; 10 AA.
XX
AC AAJ01927;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1918.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 127; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 10 AA;

Query Match      5.9%; Score 7; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAAL 24
DB      4 GGVLAAL 10

RESULT 708
AAJ00662
ID AAJ00662 standard; peptide; 10 AA.
XX
AC AAJ00662;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #653.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 116; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 10 AA;

Query Match      5.9%; Score 7; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAAL 24
DB      4 GGVLAAL 10

RESULT 707
AAJ01067
ID AAJ01067 standard; peptide; 10 AA.
XX
AC AAJ01067;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1058.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX

```

```
XX
SQ Sequence 10 AA;
Query Match 5.9%; Score 7; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
Db 4 GGVLAAL 10

RESULT 709
AAO23381
ID AAO23381 standard; peptide; 10 AA.
XX
AC AAO23381;
XX
DT 06-NOV-2003 (first entry)
XX
DE HCV NS4A peptide 15 (SeqID 36) used to identify binding to NS3 protease.
XX
KW Hepatitis C virus; HCV; serine protease inhibitor; sequelae; cirrhosis;
KW chronic active hepatitis; hepatocellular carcinoma; liver fibrosis; NS3;
KW necrosis; inflammation; bile duct change; cytostatic; antiinflammatory;
KW bovine viral diarrhoea; hepatotropic; antiviral; virucide; dengue fever;
KW NS4A.
XX
OS Hepatitis C virus.
XX
PN WO2003051910-A2.
XX
PD 26-JUN-2003.
XX
PF 13-DEC-2002; 2002WO-CA001929.
XX
PR 14-DEC-2001; 2001US-0340574P.
XX
PA (JOYC/) JOYCE M.
PA (WILL/) WILLIAMS M.
PA (HIND/) HINDSGAUL O.
PA (TYRR/) TYRREL D L.
XX
PI Joyce M, Williams M, Hindegaul O, Tyrrel DL;
XX
WPI; 2003-607859/57.
XX
New peptides useful, e.g. in the treatment of or reduction of viral load
of hepatitis C virus and associated conditions, e.g. liver fibrosis,
PT necrosis, inflammation or bile duct changes.
XX
Example 3; Fig 7; 30pp; English.
XX
This invention relates to novel hepatitis C virus (HCV) protease
inhibitors. Specifically, these inhibitors are small, hydrophobic
peptides that work by affecting the activity of the HCV serine protease
NS3, or preventing NS3 activation by inhibition of its co-factor NS4A.
CC Chronic infection with HCV can lead to serious sequelae including
CC chronic active hepatitis, cirrhosis and hepatocellular carcinoma, as well
CC as HCV associated conditions including liver fibrosis, necrosis,
CC inflammation or bile duct changes. The present invention describes these
CC peptide inhibitors as virucides, and as such they can be used to inhibit
CC HCV replication and reduce the viral load. They also have hepatotropic
CC and antiinflammatory activity and can be described as cytostatic.
CC Furthermore, the antiviral peptides derived from the relevant conserved
CC NS3 or NS4A domains can be used to treat other viruses including the
CC dengue fever virus and the bovine viral diarrhoea virus. This peptide
CC sequence, peptide 15 (SeqID 36), is part of the C-terminal deletion
CC library that was used to identify the minimal domain of NS4A that is
CC required for binding to the NS3 protease, and hence identify inhibitor
CC peptides of the invention
XX
Sequence 10 AA;
Query Match 5.9%; Score 7; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCVVIVG 38
Db 1 GCVVIVG 7

RESULT 710
ADE97767
ID ADE97767 standard; peptide; 10 AA.
XX
AC ADE97767;
XX
DT 12-FEB-2004 (first entry)
XX
DE Immunogenic HLA-A2.1 binding peptide #249.
XX
KW cytostatic; anti-inflammatory; hepatotropic; virucide; anti-HIV;
KW nephrotropic; neuroprotective; antiarthritic; antirheumatic;
KW immunosuppressive; dermatological; muscular; nephrotropic; thyromimetic;
KW haemostatic; antithyroid; antianaemic; anabolic; hypertensive;
KW immunogenic peptide composition; immune response; prostate cancer;
KW hepatitis B; hepatitis C; AIDS; renal carcinoma; cervical carcinoma;
KW lymphoma; cytomegalovirus; CMV; condyloma acuminatum;
KW autoimmune associated disorder; multiple sclerosis; rheumatoid arthritis;
KW Sjogren syndrome; scleroderma; polymyositis; dermatomyositis;
KW ankylosing spondylitis; myasthenia gravis; MG; bullous pemphigoid;
KW pemphigus; glomerulonephritis; Goodpasture's syndrome;
KW autoimmune haemolytic anaemia; Hashimoto's disease; pernicious anaemia;
KW idiopathic thrombocytopenic purpura; Grave's disease; Addison's disease;
KW human leukocyte antigen A2.1; HLA A2.1;
KW immunogenic HLA-A2.1 binding peptide.
XX
OS Synthetic.
XX
US2003185822-A1.
XX
02-OCT-2003.
XX
03-APR-2002; 2002US-00116557.
XX
05-MAR-1993; 93US-00027146.
XX
04-JUN-1993; 93US-00073205.
XX
29-NOV-1993; 93US-00159184.
XX
02-DEC-1994; 94US-00349177.
XX
(GREY/) GREY H M.
PA (SETT/) SETTE A.
PA (SIDN/) SIDNEY J.
XX
Grey HM, Sette A, Sidney J;
XX
WPI; 2004-041186/04.
XX
Immunogenic peptide composition for preventing, treating or diagnosing
pathological states, e.g. prostate cancer, hepatitis B and C, Acquired
immunodeficiency Syndrome, and renal carcinoma, includes conserved
residues at specified positions.
XX
Example 11; Page 25; 38pp; English.
XX
The invention describes an immunogenic peptide composition comprising 9
residues including a first conserved residue at a second position from N-
terminus, and a second conserved residue at C-terminal position. The
inventive peptide composition is used to elicit an immune response
against a desired antigen for preventing, treating or diagnosing
pathological states, e.g. prostate cancer, hepatitis B, hepatitis C,
AIDS, renal carcinoma, cervical carcinoma, lymphoma, cytomegalovirus
(CMV), and condyloma acuminatum. It is also used to treat autoimmune
```

CC associated disorders, e.g. multiple sclerosis, rheumatoid arthritis,
CC Sjogren syndrome, scleroderma, polymyositis, dermatomyositis, systemic
CC lupus erythematosus, juvenile rheumatoid arthritis, ankylosing
CC spondylitis, myasthenia gravis (MG), bullous pemphigoid, pemphigus,
CC glomerulonephritis, Goodpasture's syndrome, autoimmune hemolytic anemia,
CC Hashimoto's disease, pernicious anaemia, idiopathic thrombocytopenic
CC purpura, Grave's disease, and Addison's disease. The invention defines
CC positions within a motif enabling the selection of the peptides, which
CC will bind efficiently to human leukocyte antigen (HLA) A2.1. This is the
CC amino acid sequence of an immunogenic HLA-A2.1 binding peptide.
XX
XX
SQ Sequence 10 AA;

Query Match 5.9%; Score 7; DB 8; Length 10;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
|||||
Db 4 GGVLAAL 10

RESULT 711
AAJ01079
ID AAJ01079 standard; peptide; 11 AA.

XX AC AAJ01079;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #1070.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccine against hepatitis C virus.

XX PS Disclosure; Page 127; 214pp; English.

XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX
SQ Sequence 11 AA;

Query Match 5.9%; Score 7; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
|||||
Db 5 GGVLAAL 11

RESULT 712

AAJ02023

XX ID AAJ02023 standard; peptide; 11 AA.

XX AC AAJ02023;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #2014.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMUNE INC.

XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Celis E, Kubo RT, Grey HM;

XX DR WPI; 2001-308046/32.

XX PT A new composition useful as a vaccine against hepatitis C virus.

XX PS Disclosure; Page 150; 214pp; English.

XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
XX
SQ Sequence 11 AA;

Query Match 5.9%; Score 7; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
|||||
Db 5 GGVLAAL 11

RESULT 713

AAJ00694

XX ID AAJ00694 standard; peptide; 11 AA.

XX AC AAJ00694;

XX DT 02-JUL-2001 (first entry)

XX DE Hepatitis C virus epitope #685.

XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.

XX OS Hepatitis C virus.

XX PN WO200121189-A1.

XX PD 29-MAR-2001.

XX PF 19-JUL-2000; 2000WO-US019774.

XX PR 19-JUL-1999; 99US-00357737.

XX PA (EPIM-) EPIMUNE INC.
 XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
 XX PI Baker DM, Celis E, Kubo RT, Grey HM;
 XX XX WPI; 2001-308046/32.
 XX PI A new composition useful as a vaccines against hepatitis C virus.
 XX PS Disclosure; Page 117; 214pp; English.
 XX CC The present invention describes a composition comprising a prepared
 CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
 CC These are derived from HCV HLA-binding motifs. They are useful in
 CC vaccines for the prevention and treatment of HCV infection in humans. The
 CC present sequence is an epitope used in the disclosure of the invention
 XX SQ Sequence 11 AA;
 Query Match 5.9%; Score 7; DB 4; Length 11;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAAL 24
 DB 5 GGVLAAL 11
 RESULT 714
 AA023382
 ID AA023382 standard; peptide; 11 AA.
 XX AC AA023382;
 XX DT 06-NOV-2003 (first entry)
 XX DE HCV NS4A peptide 16 (SeqID 22) used to identify binding to NS3 protease.
 XX KW Hepatitis C virus; HCV; serine protease inhibitor; sequelae; cirrhosis;
 KW chronic active hepatitis; hepatocellular carcinoma; liver fibrosis; NS3;
 KW necrosis; inflammation; bile duct change; cytostatic; antiinflammatory;
 KW bovine viral diarrhoea; hepatotropic; antiviral; virucide; dengue fever;
 KW NS4A.
 XX OS Hepatitis C virus.
 XX PN WO2003051910-A2.
 XX PD 26-JUN-2003.
 XX PF 13-DEC-2002; 2002WO-CA001929.
 XX PR 14-DEC-2001; 2001US-0340574P.
 XX XX (JOYC/) JOYCE M.
 PA (WILL/) WILLIAMS M.
 PA (HIND/) HINDSGAUL O.
 XX (TYRR/) TYRREL D L.
 PI Joyce M, Williams M, Hindsgaul O, Tyrrel DL;
 XX WPI; 2003-607859/57.
 XX New peptides useful, e.g. in the treatment of or reduction of viral load
 PT of hepatitis C virus and associated conditions, e.g. liver fibrosis,
 PT necrosis, inflammation or bile duct changes.
 XX Example 3; Fig 7; 30pp; English.
 XX This invention relates to novel hepatitis C virus (HCV) protease
 CC inhibitors. Specifically, these inhibitors are small, hydrophobic
 CC peptides that work by affecting the activity of the HCV serine protease

CC NS3, or preventing NS3 activation by inhibition of its co-factor NS4A.
 CC Chronic infection with HCV can lead to serious sequelae including
 CC chronic active hepatitis, cirrhosis and hepatocellular carcinoma, as well
 CC as HCV associated conditions including liver fibrosis, necrosis,
 CC inflammation or bile duct changes. The present invention describes these
 CC peptide inhibitors as virucides, and as such they can be used to inhibit
 CC HCV replication and reduce the viral load. They also have hepatotropic
 CC and antiinflammatory activity and can be described as cytostatic.
 CC Furthermore, the antiviral peptides derived from the relevant conserved
 CC NS3 or NS4A domains can be used to treat other viruses including the
 CC dengue fever virus and the bovine viral diarrhoea virus. This peptide
 CC sequence, peptide 16 (SeqID 22), is part of the C-terminal deletion
 CC library that was used to identify the minimal domain of NS4A that is
 CC required for binding to the NS3 protease, and hence identify inhibitor
 CC peptides of the invention
 XX SQ Sequence 11 AA;
 Query Match 5.9%; Score 7; DB 6; Length 11;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 GCVVIVG 38
 DB 1 GCVVIVG 7
 RESULT 715
 AAB96865
 ID AAB96865 standard; peptide; 13 AA.
 XX AC AAB96865;
 XX DT 06-JUL-2001 (first entry)
 XX DE Hepatitis C virus NS2/3 cleavage inhibitory peptide SEQ ID NO: 23.
 XX KW Hepatitis C virus; HCV; NS2; NS3; inhibitory peptide; cleavage;
 KW replication inhibition; chimpanzee; human; infection; gene therapy.
 XX OS Hepatitis C virus.
 XX PN WO200116379-A1.
 XX PD 08-MAR-2001.
 XX PF 25-AUG-2000; 2000WO-US023444.
 XX PR 30-AUG-1999; 99US-0151395P.
 XX PA (MERI) MERCK & CO INC.
 XX PI Darke PL, Jacobs AR, Kuo LC;
 XX WPI; 2001-343059/36.
 XX Inhibiting hepatitis C virus (HCV) replication in HCV infected cell, or
 PT in a patient or treating a patient for HCV infection comprises inhibiting
 PT autocleavage of NS2/3.
 XX Disclosure; Page 47; 50pp; English.
 XX The present invention describes methods and compositions capable of
 CC preventing the replication of hepatitis C virus (HCV), involving
 CC administering a compound which inhibits NS2/3 autocleavage. Also provided
 CC are peptides capable of inhibiting this cleavage step, of which this
 CC sequence is an example. These are useful in the treatment of HCV
 CC infection in humans and chimpanzees, and in research applications, for
 CC example in studying the stabilisation of NS2/3, the effects of NS2/3 on
 CC HCV polyprotein processing and the effects of inhibiting NS2/3
 CC autocleavage
 XX SQ Sequence 13 AA;

Query Match 5.9%; Score 7; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
 Db 1 GCVVIVG 7

RESULT 716
 AAB96863
 ID AAB96863 standard; peptide; 13 AA.
 XX
 AC AAB96863;
 XX
 DT 06-JUL-2001 (first entry)
 XX
 DE Hepatitis C virus NS2/3 cleavage inhibitory peptide SEQ ID NO: 21.
 XX
 KW Hepatitis C virus; HCV; NS2; NS3; inhibitory peptide; cleavage;
 KW replication inhibition; chimpanzee; human; infection; gene therapy.
 XX
 OS Hepatitis C virus.
 XX
 FN WO200116379-A1.
 XX
 PD 08-MAR-2001.
 XX
 PP 25-AUG-2000; 2000WO-US023444.
 XX
 PR 30-AUG-1999; 99US-0151395P.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Darke PL, Jacobs AR, Kuo LC;
 XX
 DR WPI; 2001-343059/36.
 XX
 PT Inhibiting hepatitis C virus (HCV) replication in HCV infected cell, or
 PT in a patient or treating a patient for HCV infection comprises inhibiting
 PT autocleavage of NS2/3.
 XX
 PS Disclosure; Page 46; 50pp; English.
 XX
 CC The present invention describes methods and compositions capable of
 CC preventing the replication of hepatitis C virus (HCV), involving
 CC administering a compound which inhibits NS2/3 autocleavage. Also provided
 CC are peptides capable of inhibiting this cleavage step, of which this
 CC sequence is an example. These are useful in the treatment of HCV
 CC infection in humans and chimpanzees, and in research applications, for
 CC example in studying the stabilisation of NS2/3, the effects of NS2/3 on
 CC HCV polyprotein processing and the effects of inhibiting NS2/3
 CC autocleavage
 XX
 SQ Sequence 13 AA;

Query Match 5.9%; Score 7; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
 Db 1 GCVVIVG 7

RESULT 717
 AAW09243
 ID AAW09243 standard; peptide; 14 AA.
 XX
 AC AAW09243;
 XX
 DT 29-MAR-1997 (first entry)

Query Match 5.9%; Score 7; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
 Db 1 GCVVIVG 7

RESULT 718
 ADV23004
 ID ADV23004 standard; peptide; 14 AA.
 XX
 AC ADV23004;
 XX
 DT 10-MAR-2005 (first entry)
 XX
 DE HCV H77 immunogenic peptide #245.
 XX
 KW Vaccine; virucide; antigen; autoimmune disease; infection;
 KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
 KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
 KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
 KW hodgkin's lymphoma.
 XX
 OS Hepatitis C virus strain H77.
 OS
 XX WO2004108753-A1.
 XX
 FN 16-DEC-2004.
 XX
 PD 10-JUN-2004; 2004WO-AU000775.
 XX
 PF 10-JUN-2003; 2003AU-00902875.
 XX
 PR

HCV NS4A cofactor peptide (aa21-34).
 HCV; NS4A; NS3 protease; substrate; nonstructural polyprotein; inhibitor;
 assay; liver disease; hepatocellular carcinoma; tumour.
 Hepatitis C virus.
 WO9635717-A2.
 14-NOV-1996.
 09-MAY-1996; 96WO-US006389.
 12-MAY-1995; 95US-00439747.
 (SCHE) SCHERING CORP.
 Zhang R, Murray MG, Ramanathan L;
 WPI; 1996-518617/51.
 New soluble substrates for hepatitis C virus NS3 protease - are non-
 structural polyproteins and are attached to solubilising motifs, useful
 for determining protease inhibitors.
 Disclosure; Page 56; 70pp; English.
 2 Peptides (AAW09243 and AAW09244) respectively correspond to amino acids
 21-34 and 22-34 of the NS4A cofactor of hepatitis C virus (HCV), and are
 described as NS4A active mutants. NS4A improves the ability of HCV NS3
 protease (see also AAW12963) to cleave HCV nonstructural proteins. Novel
 NS3 protease-NS4A fusions (see also AAW09236-40) have been produced that
 can be used with NS3 protease substrates (AAW12957-62) in novel high
 throughput assays to identify HCV protease inhibitors of potential
 therapeutic appln
 Sequence 14 AA;

Db 8 GCWIVG 14

RESULT 720

ADV22999

ID ADV22999 standard; peptide; 14 AA.

XX AC ADV22999;

XX DT 10-MAR-2005 (first entry)

XX DE HCV H77 immunogenic peptide #240.

XX KW Vaccine; virucide; antigen; autoimmune disease; infection; immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor; breast tumor; uterine cervix tumor; prostatic cancer; colon tumor; pancreas tumor; stomach tumor; bladder tumor; kidney tumor; hodgkin's lymphoma.

XX OS Hepatitis C virus strain H77.

XX PN WO2004108753-A1.

XX PD 16-DEC-2004.

XX PF 10-JUN-2004; 2004WO-AU0000775.

XX PR 10-JUN-2003; 2003AU-00902875.

XX PR 25-MAR-2004; 2004AU-00901589.

XX PA (UYME) UNIV MELBOURNE.

XX PI Kent SJ;

XX DR WPI; 2005-031657/03.

XX PT Use of at least one set of peptides in the preparation of a medicament for modulating an immune response, and for treating cancer or yeast, viral, bacterial, protozoal and mycoplasma infections.

XX PS Disclosure; SEQ ID NO 1419; 645pp; English.

XX CC The invention relates to the use of at least one set of peptides in the preparation of a medicament for modulating an immune response, where individual peptides of a respective set comprise different portions of an amino acid sequence corresponding to a single polypeptide of interest and display partial sequence identity or similarity to at least one other peptide of the same set of peptides (i.e. they are overlapping). Also included are an antigen-presenting cell which has been contacted with the peptides above and thus presents the peptides, a population of such antigen-presenting cells, a process for producing antigen-presenting cells for modulating an immune response to a polypeptide of interest, a method for producing antigen-specific lymphocytes, a composition comprising at least one set of the peptides (and a carrier and/or diluent), a method for modulating an immune response to a polypeptide of interest comprising administering to a patient in need at least one set of the peptides, a method for treatment and/or prophylaxis of a disease or condition associated with the presence of a polypeptide of interest and a composition of matter for modulating an immune response in a subject to a target antigen. The polypeptide of interest is also a disease- or condition-associated polypeptide that is a polypeptide produced by a pathogenic organism or a cancer, and produced by a pathogenic organism selected from yeast, viruses, bacteria, helminths, protozoans and mycoplasmas. The disease- or condition-associated polypeptide is produced by a cancer selected from melanoma, lung cancer, breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic cancer, stomach cancer, bladder cancer, kidney cancer, post transplant lymphoproliferative disease (PTLD) or Hodgkin's Lymphoma. The uncultured antigen-presenting cells or their precursors are useful in the preparation of a medicament for the treatment of a disease or condition in a subject, which disease or condition is associated with the presence or aberrant expression of a target antigen, where the antigen-presenting cells or their precursors have not been subjected to activating

CC conditions but have been contacted with an antigen that corresponds to the target antigen to express a processed or modified form of the antigen for presentation to the subject's immune system. The present sequence is one of a set of overlapping immunogenic peptides derived from a Hepatitis C virus protein.

XX SQ Sequence 14 AA;

Query Match 5.9%; Score 7; DB 9; Length 14;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24

Db 8 GGVLAAL 14

|||||

RESULT 721

ADV23003

ID ADV23003 standard; peptide; 15 AA.

XX AC ADV23003;

XX DT 10-MAR-2005 (first entry)

XX DE HCV H77 immunogenic peptide #244.

XX KW Vaccine; virucide; antigen; autoimmune disease; infection; immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor; breast tumor; uterine cervix tumor; prostatic cancer; colon tumor; pancreas tumor; stomach tumor; bladder tumor; kidney tumor; hodgkin's lymphoma.

XX OS Hepatitis C virus strain H77.

XX PN WO2004108753-A1.

XX PD 16-DEC-2004.

XX PF 10-JUN-2004; 2004WO-AU0000775.

XX PR 10-JUN-2003; 2003AU-00902875.

XX PR 25-MAR-2004; 2004AU-00901589.

XX PA (UYME) UNIV MELBOURNE.

XX PI Kent SJ;

XX DR WPI; 2005-031657/03.

XX PT Use of at least one set of peptides in the preparation of a medicament for modulating an immune response, and for treating cancer or yeast, viral, bacterial, protozoal and mycoplasma infections.

XX PS Disclosure; SEQ ID NO 1423; 645pp; English.

XX CC The invention relates to the use of at least one set of peptides in the preparation of a medicament for modulating an immune response, where individual peptides of a respective set comprise different portions of an amino acid sequence corresponding to a single polypeptide of interest and display partial sequence identity or similarity to at least one other peptide of the same set of peptides (i.e. they are overlapping). Also included are an antigen-presenting cell which has been contacted with the peptides above and thus presents the peptides, a population of such antigen-presenting cells, a process for producing antigen-presenting cells for modulating an immune response to a polypeptide of interest, a method for producing antigen-specific lymphocytes, a composition comprising at least one set of the peptides (and a carrier and/or diluent), a method for modulating an immune response to a polypeptide of interest comprising administering to a patient in need at least one set of the peptides, a method for treatment and/or prophylaxis of a disease or condition associated with the presence of a polypeptide of interest and a composition of matter for modulating an immune response in a

CC subject to a target antigen. The polypeptide of interest is also a
CC disease- or condition-associated polypeptide that is a polypeptide
CC produced by a pathogenic organism or a cancer, and produced by a
CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
CC protozoans and mycoplasmas. The disease- or condition-associated
CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
CC lymphoproliferative disease (PTLD) or Hodgkin's lymphoma. The uncultured
CC antigen-presenting cells or their precursors are useful in the
CC preparation of a medicament for the treatment of a disease or condition
CC in a subject, which disease or condition is associated with the presence
CC or aberrant expression of a target antigen, where the antigen-presenting
CC cells or their precursors have not been subjected to activating
CC conditions but have been contacted with an antigen that corresponds to
CC the target antigen to express a processed or modified form of the antigen
CC for presentation to the subject's immune system. The present sequence is
CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
CC C virus protein.

XX
SQ Sequence 15 AA;

Query Match 5.9%; Score 7; DB 9; Length 15;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
Db 5 GCVVIVG 11
|||||

RESULT 722

AAU90051
ID AAU90051 standard; peptide; 17 AA.

AC AAU90051;

DT 18-JUN-2002 (first entry)

XX Insulin/insulin-like growth factor receptor-binding peptide #2007.

XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
XX ophthalmological; insulin; receptor; gene therapy; diabetes;
XX insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
XX diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.

XX Synthetic.

XX WO200172771-A2.

XX 04-OCT-2001.

XX 29-MAR-2000; 2000WO-US008528.

XX 29-MAR-2000; 2000WO-US008528.

XX (DGIB-) DGI BIOTECHNOLOGIES LLC.
XX (NOVO) NOVO NORDISK AS.

XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;
XX Briessette R, Spetzler J, Cheng W, Ostergaard S, Mandecki WS;
XX Hansen PH, Ravera M, Hsiao K;

XX WPI; 2002-025774/03.

XX Modulating insulin activity in mammalian cells, for treating e.g.
XX diabetes and tumors, comprises using peptides that bind to insulin or
XX insulin-like growth factor receptors.

XX Disclosure; Fig 3D-1; 390pp; English.

XX The invention relates to a method of modulating insulin activity in
XX mammalian cells by administering a peptide that binds the insulin

CC receptor (IR). A composition containing a peptide, optionally expressed
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin
CC agonist are useful for treating diabetes. Also, peptides that are
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
CC receptor agonists are useful for treating neurological diseases,
CC including stroke and diabetic neuropathy. The peptides are also useful in
CC screening for compounds that bind to IR or IGF-1 receptor, potential
CC therapeutics and research reagents. AAU8034-AAU9057 represent IR and/or
CC IGF-1 receptor-binding peptides and related amino acid sequences of the
CC invention

XX
SQ Sequence 17 AA;

Query Match 5.9%; Score 7; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 1 GGVLAAL 7
|||||

RESULT 723

AAR37940
ID AAR37940 standard; protein; 18 AA.

AC AAR37940;

DT 25-MAR-2003 (revised)

DT 23-SEP-1993 (first entry)

XX HCV NS-4 type 2 region 1 (1691-1708) peptide #2.

XX Non-coding region; hepatitis C virus; blood donor; type 2; type 1; HCV;
XX NS-5; phylogeny; differentiation; NS-3; core region; type 3; PCR;
XX amplify; polymerase chain reaction; primer; NS4.

XX Synthetic.

XX WO9310239-A2.

XX 27-MAY-1993.

XX 20-NOV-1992; 92WO-GB002143.

XX 21-NOV-1991; 91GB-00024696.

XX 24-JUN-1992; 92GB-00013362.

XX (COMM-) COMMON SERVICES AGENCY.

XX Simmonds P, Chan S, Yap PL;

XX WPI; 1993-182554/22.

XX DNA encoding antigenic peptide(s) of new types of hepatitis C virus - for
XX diagnosing and treating HCV infection, screening blood samples and
XX identifying different HCV types.

XX Disclosure; Page 40; 120pp; English.

XX The sequences given in AAR37938-47 are peptides which were derived from
XX the NS-4 region of the hepatitis C virus (HCV) protein. Analysis of
XX regions of the HCV genome revealed the existence of three distinct groups
XX of HCV. Analysis of the region encompassing -255 to -62 of the 5' non
XX coding region (NCR) (see AAQ43058-75) showed a difference of 9-14% in the
XX nucleotide sequences between the three groups. Two of the groups
XX identified were similar to those of HCV variants termed type 1 and 2,
XX whilst the third appeared to represent a novel type of virus. Comparison
XX of the NS3 region (see AAR37927-30) showed a high degree of sequence
XX diversity with type 3 being phylo- genetically different to type 1 and 2.
XX The same degree different- iation was noted in the NS-5 (see AAR37923-

CC 26), core region (see AAR37931) and the NS4 region (see AAQ43106-111)
CC between type 3 and type 1 sequences. (Updated on 25-MAR-2003 to correct
CC PN field.)

SQ Sequence 18 AA;

Query Match 5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||||
Db 6 PDKEVLY 12

RESULT 724

AAW82171
ID AAW82171 standard; peptide; 18 AA.

AC AAW82171;

DT 18-FEB-1999 (first entry)

DE Fluorogenic protease indicator NS3NS4/4B peptide #4.

KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
conformation change.

OS Synthetic.

FH Key Location/Qualifiers
FT Modified-site 3

FT /label= Aib
FT /note= "alpha-aminoisobutyric acid, labelled as amino
FT acid B in the specification"

FT Modified-site 4

FT /note= "epsilon-aminocaproic acid, labelled as amino acid
FT J in the specification"

FN WO9837226-A1.

PN 27-AUG-1998.

PD 20-FEB-1998; 98WO-US003000.

PF 20-FEB-1997; 97US-00802981.

PR (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

PS WPI; 1998-467579/40.

XX New fluorogenic compositions - containing 2 fluorophores separated by a
PT peptide comprising a protease binding site, used for detecting protease
PT activity in samples.

PS Disclosure; Page 27; 90pp; English.

XX AAW82023-W82240 are peptides used in the construction of a fluorogenic
CC composition which is used for the detection of protease activity in
CC biological samples. The products can be used for the detection of
CC conformational changes in nucleic acids, oligosaccharides, polysaccharides,
CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins, a
CC molecule or polymers. In addition, attachment of a hydrophobic group to a
CC molecule can be used to enhance uptake by cells. The composition is
CC composed of P = peptide comprising a protease binding site for the
CC protease, F1, F2 peptides = fluorophores where F1 is attached to the
CC amino terminal amino acid and F2 is attached to the carboxyl terminal
CC amino acid and S1, S2 peptides = when present, are peptide spacers where
CC S1, when present, is attached to the amino terminal acid, and S2, when
XX present, is attached to the carboxyl terminal amino acid

SQ Sequence 18 AA;

Query Match 5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
|||||||
Db 6 EMECSQ 12

RESULT 725

AAG73170
ID AAG73170 standard; peptide; 18 AA.

AC AAG73170;

DT 14-AUG-2001 (first entry)

DE Protease binding site #104.

KW Protease detection; peptide cleavage; enzyme activity; fluorogenic;
KW viral infection; cancer metastasis; emphysema; arthritis; thrombosis;
KW haemophilia.

OS Synthetic.

FH Key Location/Qualifiers
FT Modified-site 3

FT /label= Aib
FT /note= "2-aminoisobutyric acid"

FN WO200118238-A1.

PN 15-MAR-2001.

PD 11-SEP-2000; 2000WO-US024882.

PF 10-SEP-1999; 99US-00394019.

PR (ONCO-) ONCOIMMUNIN INC.

PI Komoriya A, Packard BS;

PS WPI; 2001-389573/41.

XX New fluorogenic compositions whose fluorescence level increases in the
PT presence of active proteases, useful for detecting and localizing
PT protease activity in biological samples, particularly in frozen tissue
PT samples.

PS Disclosure; Page 26; 86pp; English.

XX The present invention describes fluorogenic compositions which can be
CC used for the detection of protease activity. This can be useful as an
CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
CC thrombosis and arthritis. The fluorogenic compositions comprise a
CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
CC peptide is cleaved by a protease and the fluorophores can then be
CC detected. The present sequence is one of the peptides described in the
CC exemplification of the invention

SQ Sequence 18 AA;

Query Match 5.9%; Score 7; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
|||||||
Db 6 EMECSQ 12

```

RESULT 726
AAG73173
ID AAG73173 standard; peptide; 18 AA.
XX AC AAG73173;
XX DT 14-AUG-2001 (first entry)
XX DE Protease binding site #107.
XX KW Protease detection; peptide cleavage; enzyme activity; fluorogenic;
XX KW viral infection; cancer metastasis; emphysema; arthritis; thrombosis;
XX KW haemophilia.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 3
XX PT /label= Aib
XX PT /note= "2-aminoisobutyric acid"
XX WO200118238-A1.
XX PN 15-MAR-2001.
XX PD
XX PF 11-SEP-2000; 2000WO-US024882.
XX PR 10-SEP-1999; 99US-00394019.
XX PA (ONCO-) ONCOIMMUNIN INC.
XX PI Komoriya A, Packard BS;
XX DR WPI; 2001-389573/41.
XX DT
XX PT New fluorogenic compositions whose fluorescence level increases in the
XX PT presence of active proteases, useful for detecting and localizing
XX PT protease activity in biological samples, particularly in frozen tissue
XX PT samples.
XX P8 Disclosure; Page 26; 86pp; English.
XX CC The present invention describes fluorogenic compositions which can be
XX CC used for the detection of protease activity. This can be useful as an
XX CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
XX CC thrombosis and arthritis. The fluorogenic compositions comprise a
XX CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
XX CC peptide is cleaved by a protease and the fluorophores can then be
XX CC detected. The present sequence is one of the peptides described in the
XX CC exemplification of the invention
XX SQ Sequence 18 AA;
Query Match 5.9%; Score 7; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 61 EMEECSQ 67
Db 6 EMEECSQ 12
RESULT 727
ADN88410
ID ADN88410 standard; peptide; 18 AA.
XX AC ADN88410;
XX DT 12-AUG-2004 (first entry)
XX DE Fluorogenic protease indicator peptide #106.
XX KW fluorogenic; protease detection; protease inhibitor.
XX OS
XX PN WO2004108753-A1.
XX
XX OS Synthetic.
XX PN US2004096926-A1.
XX XX 20-MAY-2004.
XX PF 04-JUN-2001; 2001US-00874350.
XX PR 20-FEB-1997; 97US-00802981.
XX PR 20-FEB-1998; 98WO-US003000.
XX PR 10-SEP-1999; 99US-00394019.
XX PR 11-SEP-2000; 2000WO-US024882.
XX PA (ONCO-) ONCOIMMUNIN INC.
XX PI Packard BS, Komoriya A;
XX DR WPI; 2004-399235/37.
XX PT Fluorogenic composition useful for detecting protease activity and test
XX PT substance modulating protease activity.
XX PS Disclosure; SEQ ID NO 106; 114pp; English.
XX CC The invention relates to a fluorogenic composition (I) for detecting the
XX CC activity of a protease. (I) is useful for detecting the activity of a
XX CC protease, which involves contacting the protease with (I), where the
XX CC activity of protease is detected in a histological section, cell culture
XX CC or tissue section. The cell suspension is derived from the biological
XX CC sample chosen from tissue, blood, urine, saliva, lymph or biopsy. The
XX CC protease activity is detected by fluorescence microscopy, fluorescence
XX CC microplate reader, absorption spectroscopy or confocal fluorescent microplate
XX CC reader. (I) is useful for delivering a molecule into a cell, and for
XX CC screening a test agent for the ability to modulate the activity of the
XX CC protease. (I) is useful for detection and localisation of protease
XX CC activity in biological samples. (I) also acts as a protease inhibitor,
XX CC thus useful as protease inhibitors. (I) enables detection of the protease
XX CC activity, and provides a high intensity fluorescent signal at a visible
XX CC wavelength when they are digested by a protease. The present sequence
XX CC represents a fluorogenic protease indicator peptide of the invention.
XX SQ Sequence 18 AA;
Query Match 5.9%; Score 7; DB 8; Length 18;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 61 EMEECSQ 67
Db 6 EMEECSQ 12
RESULT 728
ADV23007
ID ADV23007 standard; peptide; 18 AA.
XX AC ADV23007;
XX DT 10-MAR-2005 (first entry)
XX DE HCV H77 immunogenic peptide #248.
XX KW Vaccine; virucide; antigen; autoimmune disease; infection;
XX KW immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
XX KW breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
XX KW pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
XX KW hodgkin's lymphoma.
XX OS Hepatitis C virus strain H77.
XX PN WO2004108753-A1.

```

XX PD 16-DEC-2004.
XX PF 10-JUN-2004; 2004WO-AU000775.
XX PP 10-JUN-2003; 2003AU-00902875.
XX PR 25-MAR-2004; 2004AU-00901589.
XX PA (UYME) UNIV MELBOURNE.
XX PI Kent SJ;
XX PT WPI; 2005-031657/03.
XX DR Use of at least one set of peptides in the preparation of a medicament
XX PT for modulating an immune response, and for treating cancer or yeast,
XX PT viral, bacterial, protozoal and mycoplasma infections.
XX XX
XX PS Disclosure; SEQ ID NO 1427; 645pp; English.
XX CC The invention relates to the use of at least one set of peptides in the
XX CC preparation of a medicament for modulating an immune response, where
XX CC individual peptides of a respective set comprise different portions of an
XX CC amino acid sequence corresponding to a single polypeptide of interest and
XX CC display partial sequence identity or similarity to at least one other
XX CC peptide of the same set of peptides (i.e. they are overlapping). Also
XX CC included are an antigen-presenting cell which has been contacted with the
XX CC peptides above and thus presents the peptides, a population of such
XX CC antigen-presenting cells, a process for producing antigen-presenting
XX CC cells for modulating an immune response to a polypeptide of interest, a
XX CC method for producing antigen-specific lymphocytes, a composition
XX CC comprising at least one set of the peptides (and a carrier and/or
XX CC diluent), a method for administering an immune response to a polypeptide of
XX CC interest comprising administering to a patient in need at least one set
XX CC of the peptides, a method for treatment and/or prophylaxis of a disease
XX CC or condition associated with the presence of a polypeptide of interest
XX CC and a composition of matter for modulating an immune response in a
XX CC subject to a target antigen. The polypeptide of interest is also a
XX CC disease- or condition-associated polypeptide that is a polypeptide
XX CC produced by a pathogenic organism or a cancer, and produced by a
XX CC pathogenic organism selected from yeast, viruses, bacteria, helminths,
XX CC protozoans and mycoplasmas. The disease- or condition-associated
XX CC polypeptide is produced by a cancer selected from melanoma, lung cancer,
XX CC breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
XX CC cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
XX CC lymphoproliferative disease (PTLD) or Hodgkin's Lymphoma. The uncultured
XX CC antigen-presenting cells or their precursors are useful in the
XX CC preparation of a medicament for the treatment of a disease or condition
XX CC in a subject, which disease or condition is associated with the presence
XX CC or aberrant expression of a target antigen, where the antigen-presenting
XX CC cells or their precursors have not been subjected to activating
XX CC conditions but have been contacted with an antigen that corresponds to
XX CC the target antigen to express a processed or modified form of the antigen
XX CC for presentation to the subject's immune system. The present sequence is
XX CC one of a set of overlapping immunogenic peptides derived from a Hepatitis
XX C virus protein.
SQ Sequence 18 AA;
Query Match 5.9%; Score 7; DB 9; Length 18;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECS 66
| | | | |
Db 12 DEMEECS 18
RESULT 729
ADV23015
ID ADV23015 standard; peptide; 18 AA.
XX
AC ADV23015;

XX 10-MAR-2005 (first entry)
XX HCV H77 immunogenic peptide #256.
XX
XX Vaccine; virucide; antigen; autoimmune disease; infection;
XX immune modulation; cancer; neoplasm; cytostatic; melanoma; lung tumor;
XX breast tumor; uterine cervix tumor; prostatic cancer; colon tumor;
XX pancreas tumor; stomach tumor; bladder tumor; kidney tumor;
XX hodgekin's lymphoma.
XX
XX Hepatitis C virus strain H77.
XX WO2004108753-A1.
XX 16-DEC-2004.
XX
XX 10-JUN-2004; 2004WO-AU000775.
XX
XX 10-JUN-2003; 2003AU-00902875.
XX 25-MAR-2004; 2004AU-00901589.
XX (UYME) UNIV MELBOURNE.
XX Kent SJ;
XX WPI; 2005-031657/03.
XX Use of at least one set of peptides in the preparation of a medicament
XX for modulating an immune response, and for treating cancer or yeast,
XX viral, bacterial, protozoal and mycoplasma infections.
XX Disclosure; SEQ ID NO 1435; 645pp; English.
XX
XX The invention relates to the use of at least one set of peptides in the
XX preparation of a medicament for modulating an immune response, where
XX individual peptides of a respective set comprise different portions of an
XX amino acid sequence corresponding to a single polypeptide of interest and
XX display partial sequence identity or similarity to at least one other
XX peptide of the same set of peptides (i.e. they are overlapping). Also
XX included are an antigen-presenting cell which has been contacted with the
XX peptides above and thus presents the peptides, a population of such
XX antigen-presenting cells, a process for producing antigen-presenting
XX cells for modulating an immune response to a polypeptide of interest, a
XX method for producing antigen-specific lymphocytes, a composition
XX comprising at least one set of the peptides (and a carrier and/or
XX diluent), a method for administering an immune response to a polypeptide of
XX interest comprising administering to a patient in need at least one set
XX of the peptides, a method for treatment and/or prophylaxis of a disease
XX or condition associated with the presence of a polypeptide of interest
XX and a composition of matter for modulating an immune response in a
XX subject to a target antigen. The polypeptide of interest is also a
XX disease- or condition-associated polypeptide that is a polypeptide
XX produced by a pathogenic organism or a cancer, and produced by a
XX pathogenic organism selected from yeast, viruses, bacteria, helminths,
XX protozoans and mycoplasmas. The disease- or condition-associated
XX polypeptide is produced by a cancer selected from melanoma, lung cancer,
XX breast cancer, cervical cancer, prostate cancer, colon cancer, pancreatic
XX cancer, stomach cancer, bladder cancer, kidney cancer, post transplant
XX lymphoproliferative disease (PTLD) or Hodgkin's Lymphoma. The uncultured
XX antigen-presenting cells or their precursors are useful in the
XX preparation of a medicament for the treatment of a disease or condition
XX in a subject, which disease or condition is associated with the presence
XX or aberrant expression of a target antigen, where the antigen-presenting
XX cells or their precursors have not been subjected to activating
XX conditions but have been contacted with an antigen that corresponds to
XX the target antigen to express a processed or modified form of the antigen
XX for presentation to the subject's immune system. The present sequence is
XX one of a set of overlapping immunogenic peptides derived from a Hepatitis
XX C virus protein.
XX
XX Sequence 18 AA;

QY 106 TYNWQKLE 112
Db |||||
4 TYNWQKLE 10

RESULT 732
AAW82174
ID AAW82174 standard; peptide; 19 AA.
XX AC
XX AAW82174;
XX DT 18-FEB-1999 (first entry)
XX DE Fluorogenic protease indicator NS3NS4A/4B peptide #7.
XX KW Protease activity; fluorophore; detection; fluorogenic; cellular uptake;
XX KW conformation change.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 3 /label= Aib
FT /note= "alpha-aminoisobutyric acid, labelled as amino
FT acid B in the specification"
FT Modified-site 4
FT /note= "epsilon-aminocaproic acid, labelled as amino acid
FT J in the specification"
XX W09837226-A1.
XX PN
XX DT 27-AUG-1998.
XX PF 20-FEB-1998; 98WO-US003000.
XX PR 20-FEB-1997; 97US-00802981.
XX PA (ONCO-) ONCOIMMUNIN INC.
XX PI Komoriya A, Packard BS;
XX WPI; 1998-467579/40.
XX
XX New fluorogenic compositions - containing 2 fluorophores separated by a
PT peptide comprising a protease binding site, used for detecting protease
PT activity in samples.
XX Disclosure; Page 27; 90pp; English.
XX
XX AAW82023-W82240 are peptides used in the construction of a fluorogenic
CC composition which is used for the detection of protease activity in
CC biological samples. The products can be used for the detection of
CC conformational changes in nucleic acids, oligosaccharides, polysaccharides,
CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,
CC steroids or polymers. In addition, attachment of a hydrophobic group to a
CC molecule can be used to enhance uptake by cells. The composition is
CC composed of P = peptide comprising a protease binding site for the
CC protease, F1, F2 peptides = fluorophores where F1 is attached to the
CC amino terminal amino acid and F2 is attached to the carboxyl terminal
CC amino acid and S1, S2 peptides = when present, are peptide spacers where
CC S1, when present, is attached to the amino terminal acid, and S2, when
CC present, is attached to the carboxyl terminal amino acid
XX Sequence 19 AA;
Query Match 5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 EMECSQ 67
Db |||||
6 EMECSQ 12

RESULT 733
AAW82176
ID AAG73176 standard; peptide; 19 AA.
XX AC
XX AAG73176;
XX DT 14-AUG-2001 (first entry)
XX DE Protease binding site #110.
XX KW Protease detection; peptide cleavage; enzyme activity; fluorogenic;
XX KW viral infection; cancer metastasis; emphysema; arthritis; thrombosis;
XX KW haemophilia.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Modified-site 3 /label= Aib
FT /note= "2-aminoisobutyric acid"
XX W0200118238-A1.
XX PN
XX PD 15-MAR-2001.
XX PF 11-SEP-2000; 2000WO-US024882.
XX PR 10-SEP-1999; 99US-00394019.
XX PA (ONCO-) ONCOIMMUNIN INC.
XX PI Komoriya A, Packard BS;
XX WPI; 2001-389573/41.
XX
XX New fluorogenic compositions whose fluorescence level increases in the
PT presence of active proteases, useful for detecting and localizing
PT protease activity in biological samples, particularly in frozen tissue
PT samples.
XX Disclosure; Page 26; 86pp; English.
XX
XX The present invention describes fluorogenic compositions which can be
CC used for the detection of protease activity. This can be useful as an
CC indicator of viral infection, cancer metastasis, haemophilia, emphysema,
CC thrombosis and arthritis. The fluorogenic compositions comprise a
CC peptide, a peptide spacer and a donor and an acceptor fluorophore. The
CC peptide is cleaved by a protease and the fluorophores can then be
CC detected. The present sequence is one of the peptides described in the
CC exemplification of the invention
XX Sequence 19 AA;
Query Match 5.9%; Score 7; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 EMECSQ 67
Db |||||
6 EMECSQ 12

RESULT 734
ADN88413
ID ADN88413 standard; peptide; 19 AA.
XX AC
XX ADN88413;
XX DT 12-AUG-2004 (first entry)
XX DE Fluorogenic protease indicator peptide #109.
XX

KW fluorogenic; protease detection; protease inhibitor.

OS Synthetic.

XX US2004096926-A1.

PN 20-MAY-2004.

PD 04-JUN-2001; 2001US-00874350.

XX 20-FEB-1997; 97US-00802981.

XX 20-FEB-1998; 98WO-US003000.

PR 10-SEP-1999; 99US-00394019.

PR 11-SEP-2000; 2000WO-US024882.

XX (ONCO-) ONCOIMMUNIN INC.

XX Packard BS, Komoriya A;

DR WPI; 2004-399235/37.

XX Fluorogenic composition useful for detecting protease activity and test substance modulating protease activity.

XX Disclosure; SEQ ID NO 109; 114pp; English.

CC The invention relates to a fluorogenic composition (I) for detecting the activity of a protease. (I) is useful for detecting the activity of a protease, which involves contacting the protease with (I), where the activity of protease is detected in a histological section, cell culture or tissue section. The cell suspension is derived from the biological sample chosen from tissue, blood, urine, saliva, lymph or biopsy. The protease activity is detected by fluorescence microscopy, fluorescence microplate reader, absorption microplate reader, flow cytometry, fluorometry, absorption spectroscopy or confocal fluorescent microplate reader. (I) is useful for delivering a molecule into a cell, and for screening a test agent for the ability to modulate the activity of the protease. (I) is useful for detection and localisation of protease activity in biological samples. (I) also acts as a protease inhibitor, thus useful as protease inhibitors. (I) enables detection of the protease activity, and provides a high intensity fluorescent signal at a visible wavelength when they are digested by a protease. The present sequence represents a fluorogenic protease indicator peptide of the invention.

XX Sequence 19 AA;

Query Match 5.9%; Score 7; DB 8; Length 19;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMERCSQ 67

Db 6 EMERCSQ 12

RESULT 735

AAAR74610

ID AAAR74610 standard; peptide; 20 AA.

XX AAAR74610;

XX 05-JAN-1996 (first entry)

DE Hepatitis C virus antigenic peptide ckk-n6.

XX Hepatitis C virus; HCV; non-A non-B; antigenic peptide; serotyping;

KW diagnosis; ckk-n6.

XX Hepatitis C virus.

OS WO9511918-A1.

PN 04-MAY-1995.

XX 28-OCT-1994; 94WO-JF001823.

XX 29-OCT-1993; 93JP-00272864.

PR 31-AUG-1994; 94JP-00207695.

XX (SRLS-) SRL INC.

XX Kumazawa T;

XX WPI; 1995-178822/23.

XX Peptide antigen recognised by hepatitis C antibodies - is used for serotyping of HCV strains in diagnosis of HCV infection.

XX Claim 2; Page 31; 75pp; Japanese.

XX AAR74609-R74618 are hepatitis C virus (HCV) antigenic peptides, which correspond to parts of the NS4, NS5 and core domains of HCV. They can be used to diagnose HCV infections, and by using two or more ligands specific to different HCV serotypes, the serotype of the HCV corresponding to the antibodies in a specimen can be determined

XX Sequence 20 AA;

Query Match 5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56

Db 9 PDKEVLY 15

RESULT 736

AAAR82177

ID AAAR82177 standard; peptide; 20 AA.

XX AAAR82177;

XX 18-FEB-1999 (first entry)

XX Fluorogenic protease indicator NS3NS4A/4B peptide #10.

DE Protease activity; fluorophore; detection; fluorogenic; cellular uptake; conformation change.

XX Synthetic.

XX Key Modified-site 3 Location/Qualifiers

FT /label= Aib
FT /note= "alpha-aminoisobutyric acid, labelled as amino acid B in the specification"

FT Modified-site 4
FT /note= "epsilon-aminocaproic acid, labelled as amino acid J in the specification"

FT Modified-site 16
FT /note= "epsilon-aminocaproic acid, labelled as amino acid J in the specification"

XX WO9837226-A1.

XX 27-AUG-1998.

XX 20-FEB-1998; 98WO-US003000.

XX 20-FEB-1997; 97US-00802981.

XX (ONCO-) ONCOIMMUNIN INC.

XX Komoriya A, Packard BS;

DR WPI; 1998-467579/40.

XX New fluorogenic compositions - containing 2 fluorophores separated by a

PT peptide comprising a protease binding site, used for detecting protease

XX activity in samples.

XX Disclosure; Page 27; 90pp; English.

XX

CC AA82023-W82240 are peptides used in the construction of a fluorogenic

CC composition which is used for the detection of protease activity in

CC biological samples. The products can be used for the detection of

CC conformational changes in nucleic acids, oligosaccharides, polysaccharides,

CC proteins, peptides, lipids, phospholipids, glycolipids, glycoproteins,

CC steroids or polymers. In addition, attachment of a hydrophobic group to a

CC molecule can be used to enhance uptake by cells. The composition is

CC composed of P = peptide comprising a protease binding site for the

CC protease, F1, F2 peptides = fluorophores where F1 is attached to the

CC amino terminal amino acid and F2 is attached to the carboxyl terminal

CC amino acid and S1, S2 peptides = when present, are peptide spacers where

CC S1, when present, is attached to the amino terminal acid, and S2, when

CC present, is attached to the carboxyl terminal amino acid

XX

SQ Sequence 20 AA;

Query Match 5.9%; Score 7; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 43;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMERCSQ 67

DB 6 EMERCSQ 12

RESULT 737

ID ADN88416 standard; peptide; 20 AA.

XX

AC ADN88416;

XX

DT 12-AUG-2004 (first entry)

XX

DE Fluorogenic protease indicator peptide #112.

XX

KW fluorogenic; protease detection; protease inhibitor.

XX

OS Synthetic.

XX

PN US2004096926-A1.

XX

PD 20-MAY-2004.

XX

PF 04-JUN-2001; 2001US-00874350.

XX

PR 20-FEB-1997; 97US-00802981.

XX

PR 10-FEB-1998; 98WO-US003000.

XX

PR 10-SEP-1999; 99US-00394019.

XX

PR 11-SEP-2000; 2000WO-US024882.

XX

PA (ONCO-) ONCOIMMUNIN INC.

XX

PI Packard BS, Komoriya A;

XX

DR WPI; 2004-399235/37.

XX

PT Fluorogenic composition useful for detecting protease activity and test

PT substance modulating protease activity.

XX

PS Disclosure; SEQ ID NO 112; 114pp; English.

XX

XX The invention relates to a fluorogenic composition (I) for detecting the

CC activity of a protease. (I) is useful for detecting the activity of a

CC protease, which involves contacting the protease with (I), where the

CC activity of protease is detected in a histological section, cell culture

CC or tissue section. The cell suspension is derived from the biological

CC sample chosen from tissue, blood, urine, saliva, lymph or biopsy. The

CC protease activity is detected by fluorescence microscopy, fluorescence

CC microplate reader, absorption microplate reader, flow cytometry,

CC fluorometry, absorption spectroscopy or confocal fluorescent microplate

CC reader. (I) is useful for delivering a molecule into a cell, and for

CC screening a test agent for the ability to modulate the activity of the

CC protease. (I) is useful for detection and localisation of protease

CC activity in biological samples. (I) also acts as a protease inhibitor,

CC thus useful as protease inhibitors. (I) enables detection of the protease

CC activity, and provides a high intensity fluorescent signal at a visible

CC wavelength when they are digested by a protease. The present sequence

CC represents a fluorogenic protease indicator peptide of the invention.

XX

SQ Sequence 20 AA;

Query Match 5.9%; Score 7; DB 8; Length 20;

Best Local Similarity 100.0%; Pred. No. 43;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMERCSQ 67

DB 6 EMERCSQ 12

RESULT 738

AAAR41149

ID AAAR41149 standard; peptide; 22 AA.

XX

AC AAAR41149;

XX

DT 25-MAR-2003 (revised)

XX

DT 22-MAR-1994 (first entry)

XX

DE HCV (type 2) peptide VIII-2 or NS4-1 (2).

XX

KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;

XX non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;

XX epitope; antibody; biotin; diagnosis; detection; vaccine.

XX

OS Synthetic.

XX

FH Key

XX

FT Modified-site 1

XX

FT Location/Qualifiers

XX

FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B=

FT biotin; X= biotinylation cpd. incorporated during

FT synthesis; Y= bond or linking gp(s). which minimises

FT steric hindrance, where Y is not a bond it is pref. 1-10

FT residues of (same or different) glycine, beta-alanine, 4-

FT aminobutyric acid, 5-aminovaleic acid or 6-aminohexanoic

FT acid; parenthesis around B and X indicate opt. presence

FT at the specified positions but B or X must be present in

FT at least one of the positions shown, B interacts with the

FT peptide to give a cpd. with greater diagnostic

FT sensitivity; A (optional)= one or more amino acids, NH2

FT or gp. which modifies the N-terminus; Z= one or more

FT amino acids, OH, NH2, or a linkage involving either of

FT these 2 gps."

XX

FT Modified-site 22

XX

FT /note= "the C-terminal comprises Y-(X)-Z"

XX

XX WO9318054-A2.

XX

PD 16-SEP-1993.

XX

XX 08-MAR-1993; 93WO-EP000517.

XX

XX 06-MAR-1992; 92EP-00400598.

XX

XX (INNO-) INNOGENETICS NV SA.

XX

XX De Leys R;

XX

```

DR WPI; 1993-303397/38.
XX
PT New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
PT increased antigenicity, useful in antibodies detection and vaccines
PT against hepatitis C, HIV and HTLV.
XX
PS Claim 4; Page 90-98; 133pp; English.
XX
CC Peptide compens. comprise at least one and pref. a combination of two,
CC three, four or more biotinylated peptides chosen from the sequences given
CC in AAR41058-R41166. The peptides represent immunologically important
CC regions of viral proteins and are prepd. by solid phase peptide
CC synthesis. The compens. are useful for the detection of antibodies to
CC HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 22 AA;
Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 10 PDKEVLY 16
|||||

RESULT 739
AAR41150
ID AAR41150 standard; peptide; 22 AA.
XX
AC AAR41150;
XX
DT 25-MAR-2003 (revised)
DT 22-MAR-1994 (first entry)
XX
DE HCV (type 2) peptide IX-2.
XX
KW Human immunodeficiency virus; HIV; hepatitis C virus; HCV;
KW non-A non-B hepatitis; NANBH; human T-cell lymphotropic virus; HTLV;
KW epitope; antibody; biotin; diagnosis; detection; vaccine.
XX
OS Synthetic.
XX
Key Location/Qualifiers
FH Modified-site 1
FT /note= "the N-terminal comprises (A)-(B)-(X)-Y; where B=
FT biotin; X= biotinylation cpd. incorporated during
FT synthesis; Y= bond or linking gp(s). which minimises
FT steric hindrance, where Y is not a bond it is pref. 1-10
FT residues of (same or different) glycine, beta-alanine, 4-
FT aminobutyric acid, 5-aminovaleic acid or 6-aminohexanoic
FT acid; parenthesis around B and X indicate opt. presence
FT at the specified positions but B or X must be present in
FT at least one of the positions shown, B interacts with the
FT peptide to give a cpd. with greater diagnostic
FT sensitivity; A (optional)= one or more amino acids, NH2
FT or gp. which modifies the N-terminus; Z= one or more
FT amino acids, OH, NH2, or a linkage involving either of
FT these 2 Gps."
FT
FT Modified-site 22
FT /note= "the C-terminal comprises Y-(X)-Z"
FT
XX
XX WO9318054-A2.
XX
XX 16-SEP-1993.
XX
XX 08-MAR-1993; 93WO-EP000517.
XX
XX 06-MAR-1992; 92EP-00400598.
XX
XX (INNO-) INNOGENETICS NV SA.
XX
PI De Leys R;
XX
DR WPI; 1993-303397/38.
XX
PT New biotinylated peptide(s) corresp. to immuno-dominant epitope(s) - with
PT increased antigenicity, useful in antibodies detection and vaccines
PT against hepatitis C, HIV and HTLV.
XX
PS Claim 4; Page 90-98; 133pp; English.
XX
CC Peptide compens. comprise at least one and pref. a combination of two,
CC three, four or more biotinylated peptides chosen from the sequences given
CC in AAR41058-R41166. The peptides represent immunologically important
CC regions of viral proteins and are prepd. by solid phase peptide
CC synthesis. The compens. are useful for the detection of antibodies to
CC HCV, and/or HIV, and/or HTLV-I or II. (Updated on 25-MAR-2003 to correct
CC PN field.)
XX
SQ Sequence 22 AA;
Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 4 PDKEVLY 10
|||||

RESULT 740
AAR99902
ID AAR99902 standard; peptide; 22 AA.
XX
AC AAR99902;
XX
DT 16-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 23-APR-1997 (first entry)
XX
DE Hepatitis C Virus antigen peptide 12.
XX
KW HCV; immunologically reactive; antigen; epitope; screen; typing;
KW antibody; immobilised.
XX
OS Hepatitis C virus; Virus.
XX
EP7266463-A2.
XX
PD 14-AUG-1996.
XX
PF 09-FEB-1996; 96EP-00101925.
XX
PR 09-FEB-1995; 95DE-01004302.
XX
PA (BOEF ) BOEHRINGER MANNHEIM GMBH.
PA (HOFF ) ROCHE DIAGNOSTICS GMBH.
XX
PI Seidel C, Wienhues-Thelen U, Schmitt U;
XX
XX WPI; 1996-364504/37.
XX
PT Antibody typing by sequential reaction with immobilised antigens - esp.
PT using new hepatitis C virus peptide(s) as antigens.
XX
XX Claim 9; Page 14; 22pp; German.
XX
CC AAR99891-920 comprise at least one immunologically active region of
CC hepatitis C virus (HCV), selected from amino acid sequences 384-414, 1738
CC -1759, 2217-2236, 2402-2419 and 2345-2357. The peptides are used in a
CC claimed method for typing antibodies in a liq. sample. Assays performed
CC using the peptides as antigens, which corresp. to regions of high
CC immunogenicity and variability, can be performed routinely using simple
CC equipment to provide accurate HCV typing. AAR99901-06 are derived from

```

CC amino acids 1738-1759 of the NS4-region of HCV. (Updated on 25-MAR-2003
 CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 16-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 22 AA;

Query Match 5.9%; Score 7; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 TNWQKLE 112
 |||||
 Db 15 TNWQKLE 21

RESULT 741

AAR99901
 ID AAR99901 standard; peptide; 22 AA.

XX
 AC AAR99901;

XX 16-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 23-APR-1997 (first entry)

XX Hepatitis C Virus antigen peptide 11.

XX HCV; immunologically reactive; antigen; epitope; screen; typing;
 KW antibody; immobilised.

XX Hepatitis C virus; Virus.

XX EP726463-A2.

XX 14-AUG-1996.

XX 09-FEB-1996; 96EP-00101925.

XX 09-FEB-1995; 95DE-01004302.

XX (BOFF) BOEHRINGER MANNHEIM GMBH.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX Seidel C, Wienhues-Thelen U, Schmitt U;

XX WPI; 1996-364504/37.

XX Antibody typing by sequential reaction with immobilised antigens - esp.
 XX using new hepatitis C virus peptide(s) as antigens.

XX Claim 9; Page 13; 22pp; German.

XX AAR99891-920 comprise at least one immunologically active region of
 CC hepatitis C virus (HCV), selected from amino acid sequences 384-414, 1738
 CC -1759, 2217-2236, 2402-2419 and 2345-2357. The peptides are used in a
 CC claimed method for typing antibodies in a liq. sample. Assays performed
 CC using the peptides as antigens, which corresp. to regions of high
 CC immunogenicity and variability, can be performed routinely using simple
 CC equipment to provide accurate HCV typing. AAR99901-06 are derived from
 CC amino acids 1738-1759 of the NS4-region of HCV. (Updated on 25-MAR-2003
 CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 22 AA;

Query Match 5.9%; Score 7; DB 2; Length 22;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 TNWQKLE 112
 |||||
 Db 15 TNWQKLE 21

RESULT 742

ADN00916

XX ID ADN00916 standard; peptide; 22 AA.

XX AC ADN00916;

XX 17-JUN-2004 (first entry)

XX Hepatitis C virus NS4 region peptide SEQ ID 11.

XX antibody typing; pathogen; autoantigen; viral strain; infection;
 KW virulence; interferon therapy response; diagnosis.

XX Hepatitis C virus.

XX EP1394548-A2.

XX 03-MAR-2004.

XX 09-FEB-1996; 2003EP-00014303.

XX 09-FEB-1995; 95DE-01004302.

XX 09-FEB-1996; 96EP-00101925.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

XX Seidel C, Wienhues-Thelen U, Schmitt U;

XX WPI; 2004-249299/24.

XX New peptides from hepatitis C virus, useful for typing virus-specific
 PT antibodies to identify the infecting strain, are derived from immunogenic
 PT viral regions.

XX Claim 3; SEQ ID NO 11; 21pp; German.

XX This invention describes a novel peptide containing at least one
 CC immunologically active region from hepatitis C virus (HCV). The method
 CC also describes a general fractional immunosorption method for typing
 CC antibodies, directed against pathogens or autoantigens. The peptides are
 CC derived from HCV of types 1a, 1b, 2a and 2b. The peptides are used to
 CC determine antibodies against HCV, particularly for typing antibodies and
 CC thus the viral strain causing infection, e.g. to assess virulence and
 CC likely response to interferon therapy. They can also be used for
 CC diagnosis of HCV infection. The invention provides a typing method
 CC suitable for routine use, without need for complex apparatus and
 CC differentiates reliably between different antibody types. Particularly,
 CC compared with known typing methods, this method saves time, sample and
 CC reagents, since it does not require preincubation of the sample with type
 CC -heterologous peptides. Since the peptides are derived from HCV regions
 CC that are simultaneously highly immunogenic and variable, they provide
 CC better typing than known peptides.

XX Sequence 22 AA;

Query Match 5.9%; Score 7; DB 8; Length 22;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 106 TNWQKLE 112
 |||||
 Db 15 TNWQKLE 21

RESULT 743

ADN00917

XX ID ADN00917 standard; peptide; 22 AA.

XX AC ADN00917;

XX 17-JUN-2004 (first entry)

XX

DE Hepatitis C virus NS4 region peptide SEQ ID 12.
XX antibody typing; pathogen; autoantigen; viral strain; infection;
KW virulence; interferon therapy response; diagnosis.
XX
OS Hepatitis C virus.
XX
PN EP1394548-A2.
XX
PD 03-MAR-2004.
XX
XX 09-FEB-1996; 2003EP-00014303.
XX
PR 09-FEB-1995; 95DE-01004302.
PR 09-FEB-1996; 96EP-00101925.
XX
XX (HOFF) ROCHE DIAGNOSTICS GMBH.
XX
XX Seidel C, Wienhues-Thelen U, Schmitt U;
XX WPI; 2004-249299/24.
XX
XX New peptides from hepatitis C virus, useful for typing virus-specific
PT antibodies to identify the infecting strain, are derived from immunogenic
PT viral regions.
XX
XX Claim 3; SEQ ID NO 12; 21pp; German.
XX
XX This invention describes a novel peptide containing at least one
CC immunologically active region from hepatitis C virus (HCV). The method
CC also describes a general fractional immunosorption method for typing
CC antibodies, directed against pathogens or autoantigens. The peptides are
CC derived from HCV of types 1a, 1b, 2a and 2b. The peptides are used to
CC determine antibodies against HCV, particularly for typing antibodies and
CC thus the viral strain causing infection, e.g. to assess virulence and
CC likely response to interferon therapy. They can also be used for
CC diagnosis of HCV infection. The invention provides a typing method
CC suitable for routine use, without need for complex apparatus and
CC differentiates reliably between different antibody types. Particularly,
CC compared with known typing methods, this method saves time, sample and
CC reagents, since it does not require preincubation of the sample with type
CC -heterologous peptides. Since the peptides are derived from HCV regions
CC that are simultaneously highly immunogenic and variable, they provide
CC better typing than known peptides.
XX
XX Sequence 22 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 22;
Best Local Similarity 100.0%; Pred. No. 46; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;
QY 106 TNWQKLE 112
DB 15 TNWQKLE 21
|||||
RESULT 744
AAW37388
ID AAW37388 standard; peptide; 24 AA.
AC AAW37388;
XX
XX 27-AUG-2003 (revised)
DT 11-MAR-1998 (first entry)
XX
XX Hepatitis C virus NS4-II2 protein 1690-1713.
DE
XX Hepatitis C virus; HCV; chimeric; antigen; detection; core region;
KW epitope; NS3; NS4; infection.
XX
OS Hepatitis C virus.
XX
XX JP09278794-A.
PN

XX 28-OCT-1997.
PD
XX 10-FEB-1997; 97JP-00027015.
PF
XX 09-FEB-1996; 96JP-00024045.
PR
XX (TOFU) TONEN CORP.
PA
XX WPI; 1998-022248/03.
DR
XX New chimeric peptide antigen derived from hepatitis C virus protein -
PT useful for detecting HCV infections.
PT
XX Disclosure; Page 26; 30pp; Japanese.
PS
XX The present sequence represents a Hepatitis C virus (HCV) protein
CC sequence from the disclosure of the present specification. The present
CC specification describes a chimeric HCV peptide antigen which comprises at
CC least 2 peptide epitope regions from the HCV polyprotein core region, 2
CC peptide epitope regions from the NS3 region and at least 2 peptide
CC epitope regions from the NS4 region. The antigen binds specifically with
CC an antibody produced by a human infected by HCV. The peptide can detect a
CC wide range of HCV infections with high sensitivity. (Updated on 27-AUG-
CC 2003 to correct OS field.)
XX
XX Sequence 24 AA;
SQ
Query Match 5.9%; Score 7; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 50; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;
QY 50 PDKEVLY 56
DB 11 PDKEVLY 17
|||||
RESULT 745
AAW50788
ID AAW50788 standard; peptide; 25 AA.
AC AAW50788;
XX
XX 27-AUG-2003 (revised)
DT 27-JUL-1998 (first entry)
XX
XX Peptide used in immunoassay kit for hepatitis C virus patients.
DE
XX Hepatitis C virus; immuno-assay; non-structural region 4A; NS4A;
KW non-structural region 4B; NS4B.
XX
OS Hepatitis C virus.
OS
XX JPI0019897-A.
PN
XX 23-JAN-1998.
XX
XX 27-JUN-1996; 96JP-00167497.
PF
XX 27-JUN-1996; 96JP-00167497.
PR
XX (TOFU) TONEN CORP.
PA
XX WPI; 1998-149118/14.
DR
XX New immuno-assay kit - useful for predicting effect of therapeutic on
PT hepatitis C virus patients.
PT
XX Example 1; Page 4; 9pp; Japanese.
PS
XX The invention relates to an immunoassay kit for predicting an effect of a
CC therapeutic composition, for the treatment of chronic hepatitis. The
CC method comprises: (a) a peptide having a sequence comprising at least 5

CC CONH2. These new peptides are up to 1800 times more potent than hGRF(1-
 CC 26)NH2 in stimulating growth hormone release. They do not possess Arg and
 CC Met residues at the C-terminus, thereby decreasing the cost of synthesis
 CC and enhancing their stability. They can be used: to treat growth hormone
 CC deficiency, e.g. in the elderly or in children of short stature; to
 CC stimulate tissue growth, protein metabolism, carbohydrate metabolism,
 CC lipid metabolism, mineral metabolism and connective tissue metabolism; to
 CC treat catabolic states; to stimulate immune function; to enhance natural
 CC sleep patterns; or to stimulate the growth of animals, e.g. livestock.
 CC The present sequence is one of 31 specifically claimed examples of the
 CC new variants
 XX
 SQ Sequence 26 AA;

Query Match 5.9%; Score 7; DB 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 53;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 | | | | |
 Db 13 VLAALAA 19

RESULT 748

AAW16439
 ID AAW16439 standard; peptide; 27 AA.

AC AAW16439;

XX 18-JUN-1997 (first entry)

DT Human growth factor releasing hormone analogue.

XX Human; growth; hormone; releasing factor; hGRF; analogue; release;
 KW stimulation; tissue; protein; carbohydrate; mineral; lipid; connective;
 KW metabolism; treatment; catabolic state; immune function; enhancement;
 KW natural sleep pattern; animal; livestock; stability.

XX Homo sapiens.

XX Key Location/Qualifiers

PF Misc-difference 2 /note= "D-form residue"

FT Modified-site 27 /note= "amidated"

FT WO9632126-A1.

XX 17-OCT-1996.

XX 03-APR-1996; 96WO-US004582.

XX 14-APR-1995; 95US-00421987.

PR 06-SEP-1995; 95US-00524337.

XX (TULA) TULANE EDUCATIONAL FUND.

XX Coy DH, Murphy W;

XX WPI; 1996-476839/47.

XX New growth hormone-releasing factor analogue peptide(s) - used to
 FT stimulate the release of growth hormone, e.g. for treating growth hormone
 FT deficiency disorders.

XX Claim 23; Page 32; 34pp; English.

XX The present peptide is a human growth hormone releasing factor (hGRF)
 CC analogue, which stimulates the release of growth hormone. It can be used,
 CC e.g. to stimulate growth in patients, tissue growth and protein,
 CC carbohydrate, mineral, lipid and connective tissue metabolism. It can
 CC also be used to treat catabolic states, stimulate immune function,
 CC enhance natural sleep patterns and stimulate animal growth, e.g.

CC livestock. The hGRF analogue does not possess an undesirable C-terminal
 CC Arg or Met residue, thereby decreasing its synthesis cost and enhancing
 CC its stability

SQ Sequence 27 AA;

Query Match 5.9%; Score 7; DB 2; Length 27;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 | | | | |
 Db 13 VLAALAA 19

RESULT 749

AAW52784
 ID AAW52784 standard; peptide; 27 AA.

AC AAW52784;

XX 02-MAR-1999 (first entry)

DT Human growth hormone releasing factor variant #27.

DE Human growth hormone releasing factor; hGRF; variant.

XX Homo sapiens.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 2 /note= "D-form residue"

FT Modified-site 27 /note= "Ala-NH2"

XX US5847066-A.

XX 08-DEC-1998.

XX 12-APR-1996; 96US-00631421.

XX 14-APR-1995; 95US-00421987.

PR 06-SEP-1995; 95US-00524337.

XX (TULA) TULANE EDUCATIONAL FUND.

XX Murphy W, Coy DH;

XX WPI; 1999-059146/05.

XX Peptide analogues of growth hormone-releasing factor - useful as human or
 XX animal growth promoters, etc.

XX Claim 23; Col 22; 12pp; English.

XX The invention relates to new variants of human growth hormone releasing
 CC factor (hGRF) which contain 23-28 amino acid residues and which differ
 CC from native hGRF at least at positions 8, 9, 16, 18, 24, 25, 27 and 28.
 CC More specifically, the new variants are of formula (R1)(R2)Al-A2-Arg-Ala-
 CC Ile-Phe-Thr-A8-A9-Al0-Arg-Al2- Val-Leu-Al5-Al6-Leu-Al8-Ala-Arg-A21-A22-
 CC Leu-A24-A25-A26-A27-A28-R3 where: Al is the D or L isomer of Tyr or His
 CC or is absent; A2 is Ala or the D or L isomer of Ala, N-Me-Ala or Arg; A8
 CC is Ala, Aib or Gly; A9 is Ala, Aib or Gly; A10 is Phe or pX-Phe, where X
 CC is OH, Me or halo; A12 is Lys or N-epsilon-Lys (X is 1-6C alkyl, 1-6C
 CC acyl, 1-6C hydroxyalkyl or 2-6C hydroxyacyl); A15, A16 and A18 are Ala,
 CC Aib or Gly; A21 is Lys or N-epsilon-Lys; A22 is Ala, Aib, Gly, Leu,
 CC Ile, Val, Nle, Nva or Abu; A24 is Ala, Aib, Gaba, Gly or His; A25 and A26
 CC are Ala, Aib, Gaba, Gly or His or are absent; A27 is Ala, Aib, Gly, Leu,
 CC Ile, Val, Nle, Nva, Abu, Gaba, beta-Ala, Ava or His or is absent; A28 is
 CC Aib, the D or L isomer of Ala, Gaba or His, or is absent; R1 and R2
 CC (which are N-terminal substituents) are H, 1-12C alkyl, 7-20C
 CC phenylalkyl, 11-20C naphthylalkyl, 1-12C hydroxyalkyl, 7-20C

CC hydroxyphenylalkyl, 11-20C hydroxynaphthylalkyl or COB1, where E1 is 1-
 CC 12C alkyl, 7-20C phenylalkyl, 11-20C naphthylalkyl, 1-12C hydroxyalkyl, 7-
 CC 20C hydroxyphenylalkyl or 11-20C hydroxynaphthylalkyl; and R3 (which is a
 CC substituent on the C-terminal carbonyl group) is OH, NH2, 1-12C alkoxy or
 CC NH-Y-CH2-Z, where Y is a 1-12C hydrocarbon group and Z is H, OH, COOH or
 CC CONH2. These new peptides are up to 1800 times more potent than hGRF(1-
 CC 26)NH2 in stimulating growth hormone release. They do not possess Arg and
 CC Met residues at the C-terminus, thereby decreasing the cost of synthesis
 CC and enhancing their stability. They can be used: to treat growth hormone
 CC deficiency, e.g. in the elderly or in children of short stature; to
 CC stimulate tissue growth, protein metabolism, carbohydrate metabolism,
 CC lipid metabolism, mineral metabolism and connective tissue metabolism; to
 CC treat catabolic states; to stimulate immune function; to enhance natural
 CC sleep patterns; or to stimulate the growth of animals, e.g. livestock.
 CC The present sequence is one of 31 specifically claimed examples of the
 CC new variants

SQ Sequence 27 AA;

Query Match 5.9%; Score 7; DB 2; Length 27;
 Best Local Similarity 100.0%; Pred. No. 55;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 |||||
 DB 13 VLAALAA 19

RESULT 750

ID AAW16419
 AD AAW16419 standard; peptide; 28 AA.

XX AC AAW16419;

DT 18-JUN-1997 (first entry)

DE Human growth factor releasing hormone analogue #7.

KW Human; growth; hormone; releasing factor; hGRF; analogue; release;
 KW stimulation; tissue; protein; carbohydrate; mineral; lipid; connective;
 KW metabolism; treatment; catabolic state; immune function; enhancement;
 KW natural sleep pattern; animal; livestock; stability; increased; potency.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Misc-difference 2 /note= "D-form residue"

FT Modified-site 28 /note= "amidated"

FT WO9632126-A1.

XX 17-OCT-1996.

XX 03-APR-1996; 96WO-US004582.

XX 14-APR-1995; 95US-00421987.

XX 06-SEP-1995; 95US-00524337.

XX (TULA) TULANE EDUCATIONAL FUND.

XX Coy DH, Murphy W;

XX WPT; 1996-476839/47.

XX New growth hormone-releasing factor analogue peptide(s) - used to
 FT stimulate the release of growth hormone, e.g. for treating growth hormone
 PT deficiency disorders.

XX Claim 22; Page 29; 34pp; English.

XX The present peptide is a human growth hormone releasing factor (hGRF)

CC analogue, which stimulates the release of growth hormone. It can be used,
 CC e.g. to stimulate growth in patients, tissue growth and protein.
 CC carbohydrate, mineral, lipid and connective tissue metabolism. It can
 CC also be used to treat catabolic states, stimulate immune function,
 CC enhance natural sleep patterns and stimulate animal growth, e.g.
 CC livestock. The hGRF analogue does not possess an undesirable C-terminal
 CC Arg or Met residue, thereby decreasing its synthesis cost and enhancing
 CC its stability. Also, in an assay to test its ability to release growth
 CC hormone from rat pituitary cells in vitro, it had a 560-fold greater
 CC potency than native hGRF(1-26)NH2

XX Sequence 28 AA;

Query Match 5.9%; Score 7; DB 2; Length 28;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 |||||
 DB 13 VLAALAA 19

RESULT 751

ID AAW52764
 AD AAW52764 standard; peptide; 28 AA.

XX AC AAW52764;

DT 02-MAR-1999 (first entry)

DE Human growth hormone releasing factor variant #7.

KW Human growth hormone releasing factor; hGRF; variant.

XX OS Homo sapiens.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 2 /note= "D-form residue"

FT Modified-site 28 /note= "Ala-NH2"

XX US5847066-A.

XX 08-DEC-1998.

XX 12-APR-1996; 96US-00631421.

XX 14-APR-1995; 95US-00421987.

XX 06-SEP-1995; 95US-00524337.

XX (TULA) TULANE EDUCATIONAL FUND.

XX Murphy W, Coy DH;

XX WPT; 1999-059146/05.

XX Peptide analogues of growth hormone-releasing factor - useful as human or
 FT animal growth promoters, etc.

XX Claim 22; Col 19; 12pp; English.

XX The invention relates to new variants of human growth hormone releasing
 CC factor (hGRF) which contain 23-28 amino acid residues and which differ
 CC from native hGRF at least at positions 8, 9, 16, 18, 24, 25, 27 and 28.
 CC More specifically, the new variants are of formula (R1)(R2)A1-A2-Arg-Ala-
 CC Ile-Phe-Thr-A8-A9-A10-Arg-A12- Val-Leu-A15-Leu-A18-Ala-Arg-A21-A22-
 CC Leu-A24-A25-A26-A27-A28-R3 where: A1 is the D or L isomer of Tyr or His
 CC or is absent; A2 is Ala or the D or L isomer of Ala, N-Me-Ala or Arg; A8
 CC is Ala, A1b or Gly; A9 is Ala, A1b or Gly; A10 is Phe or pX-Phe, where X
 CC is OH, Me or halo; A12 is Lys or N-epsilon-Lys (X is 1-6C alkyl, 1-6C
 CC acyl, 1-6C hydroxyalkyl or 2-6C hydroxyacyl); A15, A16 and A18 are Ala,

CC Aib or Gly; A21 is Lys or N-epsilon-X-Lys; A22 is Ala, Aib, Gly, Leu,
CC Ile, Val, Nle, Nva or Abu; A24 is Ala, Aib, Gaba, Gly or His; A25 and A26
CC are Ala, Aib, Gaba, Gly or His or are absent; A27 is Ala, Aib, Gly, Leu,
CC Ile, Val, Nle, Nva, Abu, Gaba, beta-Ala, Ava or His or is absent; A28 is
CC Aib, the D or L isomer of Ala, Gaba or His, or is absent; R1 and R2
CC (which are N-terminal substituents) are H, 1-12C alkyl, 7-20C
CC phenylalkyl, 11-20C naphthylalkyl, 1-12C hydroxyalkyl, 7-20C
CC hydroxyphenylalkyl, 11-20C hydroxynaphthylalkyl or COEt, where Et is 1-
CC 12C alkyl, 7-20C phenylalkyl, 11-20C naphthylalkyl, 1-12C hydroxyalkyl, 7-
CC 20C hydroxyphenylalkyl or 11-20C hydroxynaphthylalkyl; and R3 (which is a
CC substituent on the C-terminal carbonyl group) is OH, NH2, 1-12C alkoxy or
CC NH-Y-CH2-Z, where Y is a 1-12C hydrocarbon group and Z is H, OH, COOH or
CC CONH2. These new peptides are up to 1800 times more potent than hGRF(1-
CC 26)/NH2 in stimulating growth hormone release. They do not possess Arg and
CC Met residues at the C-terminus, thereby decreasing the cost of synthesis
CC and enhancing their stability. They can be used: to treat growth hormone
CC deficiency, e.g. in the elderly or in children of short stature; to
CC stimulate tissue growth, protein metabolism, carbohydrate metabolism,
CC lipid metabolism, mineral metabolism and connective tissue metabolism; to
CC treat catabolic states; to stimulate immune function; to enhance natural
CC sleep patterns; or to stimulate the growth of animals, e.g. livestock.
CC The present sequence is one of 31 specifically claimed examples of the
CC new variants
XX
SQ Sequence 28 AA;

Query Match 5.9%; Score 7; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
Db 13 VLALAA 19
|||||

RESULT 752
AAU84714
ID AAU84714 standard; peptide; 30 AA.
XX
AC AAU84714;
XX
DT 08-MAY-2002 (first entry)
XX
DE HCV HepC1a segment 117.
XX
KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
KW viral infection; human immunodeficiency virus; melanoma;
KW bacterial infection; Salmonella; Legionella; parasitic infection;
KW Trypanosoma; Toxoplasma; Giardia.
XX
OS Hepatitis C virus.
XX
PN WO200190197-A1.
XX
PD 29-NOV-2001.
XX
PF 25-MAY-2001; 2001WO-AU000622.
XX
PR 26-MAY-2000; 2000AU-00007761.
XX
PA (AUSU) UNIV AUSTRALIAN NAT.
XX
PI Thomson SA, Ramshaw IA;
XX
DR WPI, 2002-147575/19.
DR N-PDB; ABK36552.
XX
PT New synthetic polypeptides having several different segments of at least
PT one parent polypeptide linked together differently compared to the
PT linkage in the parent polypeptide, for inducing immune response against a
PT pathogen or cancer.
XX
PS Example 2; Fig 26; 364pp; English.

XX The invention relates to a new synthetic polypeptide (I) comprising
CC several different segments of at least one parent polypeptide linked
CC together in a different relationship relative to their linkage in the
CC parent polypeptide to impede, abrogate or otherwise alter at least one
CC function associated with the parent polypeptide and for inducing an
CC immune response against a pathogen or cancer. Also included are a
CC synthetic polynucleotide encoding and a computer system for designing the
CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
CC are referred to as a Savine. The synthetic polypeptide is useful for
CC modulating immune responses preferably directed against a pathogen or a
CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
CC oesophagus, brain, testicle, uterus), as potentiating agents.
CC Compositions comprising the polypeptide may be used in the treatment or
CC prophylaxis against viral (such as infections caused by HIV (human
CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
CC Salmonella, Streptococcal, Legionella and Mycobacterium) or parasitic
CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
CC a peptide derived from a parent protein used to construct a savine of the
CC invention
XX
SQ Sequence 30 AA;

Query Match 5.9%; Score 7; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWOKLE 112
Db 14 TNWOKLE 20
|||||

RESULT 753
AAU16755
ID AAU16755 standard; peptide; 34 AA.
XX
AC AAU16755;
XX
DT 21-JUL-1999 (first entry)
XX
DE Calcitonin peptide derivative 3.
XX
KW Calcitonin; treatment; bone disease; Paget's disease; hypercalcaemia;
KW osteoporosis.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 34
FT /note= "C-terminal amide"
XX
PN WO9629343-A1.
XX
PD 26-SEP-1996.
XX
PF 15-MAR-1996; 96WO-JP000666.
XX
PR 20-MAR-1995; 95JP-00061026.
XX
PA (KYOW) KYOWA HAKKO KOGYO KK.
XX
PI Shibata K, Yamaaki M, Hamada M, Tamaaki T, Kosaka N, Sato S;
XX WPI, 1996-443134/44.
DR
XX New calcitonin deriva. - used for the treatment and prevention of Paget's
XX disease, hypercalcaemia and osteoporosis.
XX
PS Example; Page 10; 62pp; Japanese.

XX The invention provides derivatives of calcitonin which are of the formula
 CC Cyclo(2-(X)m-Asp)-(Trp)n-Y, in which Z = Gly or Cys; X = alpha-amino acid
 CC residue; Y = natural calcitonin, natural calcitonin partial peptide or
 CC natural calcitonin related peptide; m=5-8; and n=0-3; provided that when
 CC m=5, the four residues on the C-terminal side of -(X)m- are not the same
 CC as the sequence at the 3-6 position of natural calcitonin. The
 CC derivatives are useful in the treatment and prevention of bone diseases
 CC such as Paget's disease, hypercalcaemia and osteoporosis. Sequences
 CC AAY16753-766 represent examples of calcitonin derivatives
 XX SQ Sequence 34 AA;

Query Match 5.9%; Score 7; DB 2; Length 34;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 DB 10 VLAALAA 16
 |||||

RESULT 754

AAV67801
 ID AAY67801 standard; peptide; 35 AA.

XX AC AAY67801;

XX DT 23-MAR-2000 (first entry)

XX DE Peptide #201 for detecting hepatitis C virus infection.

XX KW Hepatitis C virus; HCV; increased structural stability; NS4 region;
 diagnostic antigen.

XX OS Synthetic.

XX PN WO9962945-A2.

XX PD 09-DEC-1999.

XX PF 04-JUN-1999; 99WO-US012446.

XX PR 05-JUN-1998; 98US-0088229P.

XX PR 01-SEP-1998; 98US-0098705P.

XX PR 15-SEP-1998; 98US-0100422P.

XX PR 28-JAN-1999; 99WO-US001726.

XX (PEPT-) PEPTIDE SOLUTIONS INC.

XX Chowdhury MA, Bernstein D, Molsenbocker MA;

XX WPI; 2000-086953/07.

XX Improving properties of peptides for use as diagnostic antigens or for
 PT preventing or treating infections.

XX PS Claim 55; Page 73; 83pp; English.

XX This is a peptide related to the immunoreactive region of the NS4 region
 CC of hepatitis C virus (HCV). The peptide is useful for detecting HCV
 CC infection. The invention relates to peptides derived from HCV and also
 CC HIV-1 which have been modified for use as diagnostic antigens in the
 CC treatment or prevention of infection. The structural stability of the
 CC peptides can be increased in four different ways; through the replacement
 CC of a hydrophobic amino acid with a less hydrophobic amino acid; through
 CC an increase in the amount of secondary structure (i.e. alpha helix) in
 CC the peptide; through the removal of a positive charge from the peptide,
 CC or through the constraint of the epitopic sequence via the formation of a
 CC covalent crosslink. Modified peptides of the invention are used to detect
 CC infectious agents specifically HCV. Other detectable agents include HIV-1
 CC Group O viruses; human T-cell lymphotropic virus-I or -II; and the
 CC causative agent of syphilis. The peptides can be used for prevention or

CC treatment of infections (e.g. as vaccines, or where expressed from a
 CC transgene). More generally almost any peptide can be similarly modified,
 CC e.g. cytokines or interferons; major histocompatibility complex antigens;
 CC hormones; growth factors; tumour markers or suppressors, or antigens from
 CC many other pathogens

XX SQ Sequence 35 AA;

Query Match 5.9%; Score 7; DB 3; Length 35;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
 DB 25 PDKEVLY 31
 |||||

RESULT 755

ABB04543
 ID ABB04543 standard; peptide; 35 AA.

XX AC ABB04543;

XX DT 21-MAR-2002 (first entry)

XX DE Hepatitis C capsid protein immunogenic peptide SEQ ID NO: 23.

XX KW Hepatitis C virus; immunogenic peptide; viral protein; antiviral;

XX T cell immune response; hepatotropic; antiinflammatory; vaccine.

XX OS Hepatitis C virus.

XX PN FR2809402-A1.

XX PD 30-NOV-2001.

XX PF 26-MAY-2000; 2000FR-00006744.

XX PR 26-MAY-2000; 2000FR-00006744.

XX (SEDA-) SEDAC SOC ETUD & DEV ANTIGENES COMBINATO.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX (INSP) INST PASTEUR LILLE.

XX PI Bertrand G, Gras MH, Bouzidi A, Auriault C;

XX WPI; 2002-043465/06.

XX Preparing immunogenic peptides, useful in vaccination against hepatitis C

XX virus, by convergent combinatorial synthesis from T cell epitopes.

XX Claim 7; Fig 23; 49pp; French.

XX The present invention relates to a method for preparing immunogenic
 CC peptides that correspond to restricted regions of viral proteins
 CC implicated in a T cell immune response and to libraries of convergent
 CC combinatorial peptides corresponding to the restricted regions. The
 CC peptides can be used to prepare vaccines, particularly for use in the
 CC treatment and prevention of hepatitis C virus infection. The present
 CC sequence is an immunogenic peptide derived from the hepatitis C capsid
 CC protein described in the exemplification of the invention

XX SQ Sequence 35 AA;

Query Match 5.9%; Score 7; DB 5; Length 35;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNWQKLE 112
 DB 9 TNWQKLE 15
 |||||

CC specific to different HCV serotypes, the serotype of the HCV
XX corresponding to the antibodies in a specimen can be determined

SQ Sequence 40 AA;
Query Match 5.9%; Score 7; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
Db 11 PDKEVLY 17
RESULT 759
ID AAW01866
AAW01866 standard; peptide; 40 AA.
AC AAW01866;
XX
XX 07-NOV-1997 (first entry)
DE HCV NS-4 prototype immunoreactive peptide 2.
KW immunoreactive; HCV; Hepatitis C Virus; antibody; screening; blood;
KW diagnose; infection; immunoassay; core protein; NS-4; NS-5.
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FT Misc-difference 8
FT /note= "can be substituted with Norleucine"
FT Misc-difference 13
FT /note= "can be substituted with Ornithine"
FT Misc-difference 20
FT /note= "can be substituted with Hydroxyproline"
FT Misc-difference 28
FT /note= "can be substituted with Arginine"
FT Misc-difference 35
FT /note= "can be substituted with Ornithine"
XX DE19540105-C1.
XX
XX 20-FEB-1997.
XX
XX 27-OCT-1995; 95DE-01040105.
XX
XX 19-SEP-1995; 95US-00530550.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Hosein B, Wang CY;
XX WPI; 1997-120479/12.
XX
XX Compasn. contg. synthetic peptide(s) reactive with hepatitis C related
XX antibodies - used for screening blood and diagnosis of infection.
XX
XX Claim 9; Page 31; 46pp; German.
XX
XX This peptide is the prototype peptide for design of synthetic peptides
XX immunoreactive with Hepatitis C virus (HCV) antibodies against the NS-4
XX protein. The synthetic peptides containing substitutions as indicated in
XX the FT above, can be used in a novel composition which comprises at least
XX one linear or branched peptide (A) of formulae: P1-Y; P2-X; P4X2X;
XX P8X4X2X; or P16X8X4X2X; Y = COOH or CONH2 at the peptide C terminus; X =
XX residues of amino acids (aa), or analogues with 2 amino gps. and one COOH
XX gp., each gp. able to form a peptide bond; P = peptide that is
XX specifically immunoreactive with HCV antibodies. The compositions are
XX used to detect HCV antibodies, e.g. for screening blood products, and to
XX diagnose HCV infection by immunoassay of serum, tissue (extracts), and
XX body fluids, pref. by ELISA, sandwich or passive haemagglutination tests

SQ Sequence 40 AA;
Query Match 5.9%; Score 7; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 TNWQKLE 112
Db 24 TNWQKLE 30
RESULT 760
ABP00036
ID ABP00036 standard; protein; 68 AA.
XX
XX AC ABP00036;
XX
XX 24-JUN-2002 (first entry)
DE Human ORFX protein sequence SEQ ID NO:54.
XX
XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis.
XX
XX OS Homo sapiens.
XX
XX PN WO200192523-A2.
XX
XX PD 06-DEC-2001.
XX
XX PF 29-MAY-2001; 2001WO-US010836.
XX
XX PR 30-MAY-2000; 2000US-0206132P.
XX PR 29-AUG-2000; 2000US-0228716P.
XX
XX PA (CURA-) CURAGEN CORP.
XX
XX PI Shimkets RA, Leach MD;
XX
XX DR WPI; 2002-106308/14.
XX DR N-PSDB; ABN15788.
XX
XX PT Novel human polypeptides and polynucleotides useful for diagnosing,
XX PT preventing and treating cardiovascular disease, neurodegenerative,
XX PT hyperproliferative disorders and autoimmune disorders.
XX
XX PS Disclosure; SEQ ID NO 54; 1037pp; English.
XX
XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX in the specification). ABN15762 to ABN27252 encode the human ORFX
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX treating or preventing a pathology associated with an ORFX-associated
XX syndrome in humans, and in the manufacture of a medicament for treating a
XX disorder associated with ORFX-associated disorder. ORFX polynucleotide
XX sequences can be used in gene therapy. ORFX sequences can be used in the
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
XX psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
XX osteoarthritis, neurodegenerative disorders, disorders related to organ
XX transplantation, cardiovascular diseases, diabetes mellitus, systemic
XX lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
XX storage disease, various immune deficiencies and disorders, infectious
XX diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
XX arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
XX disease and autoimmune inflammatory eye disease. ORFX proteins are also
XX useful for treating burns, incisions, ulcers, for treating osteoporosis,
XX bone degenerative disorders, or periodontal disease, and for gut

CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage. N.B. The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 68 AA;

Query Match 5.9%; Score 7; DB 5; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 IPIVTT 106
Db 29 IPIVTT 35
|||||

RESULT 761
AAW22595
ID AAW22595 standard; protein; 75 AA.

XX AAW22595;
XX
XX
DT 12-MAR-1998 (first entry)
XX
DE Human hepatitis C virus NS4 antigen HC280.

XX Human; hepatitis C virus; antigen; recombinant protein; diagnosis;
XX immunoassay.

OS Hepatitis C virus.

XX Key Location/Qualifiers
FH Modified-site 1

FT /note= "Attached to beta-galactosidase"

FT Misc-difference 26..27
FT /note= "The present sequence is decoded by AAT80405 but
FT appears to omit 13 amino acid residues between these
FT positions"

XX RU2073718-C1.

XX
PD 20-FEB-1997.

XX
PF 03-JUN-1993; 93RU-00029440.

XX
PR 03-JUN-1993; 93RU-00029440.

XX (BIOS-) BIOSERVIS BIOTECHN CO.

XX Loparev VN, Krasnykh VN, Blinov VM;

XX WPI; 1997-423442/39.

XX N-PSDB; AAT80405.

XX DNA fragment HC280 coding for human hepatitis C virus NS4 antigen - for
XX producing recombinant polypeptide useful in immuno:diagnosis of hepatitis
XX C.

XX Claim 2; Col 12; 6pp; Russian.

XX The present sequence is encoded by the new chemically synthesised DNA
XX fragment HC280 which determines the synthesis of recombinant polypeptide
XX HC280, capable of binding antibodies to the product of the NS4 gene of
XX human hepatitis C virus (HCV). The DNA fragment has a defined sequence of
XX 283 base pairs. The NS4 antigen is useful in immunoassays for the
XX diagnosis of human hepatitis C virus

XX Sequence 75 AA;

Query Match 5.9%; Score 7; DB 2; Length 75;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
Db 6 GCVVIVG 12
|||||

RESULT 762

AAAR49654
ID AAR49654 standard; peptide; 87 AA.

XX AAR49654;

XX
DT 25-MAR-2003 (revised)

DT 02-AUG-1994 (first entry)

XX HCV peptide C14-2-2.

XX HCV; hepatitis C virus group-I; hepatitis C virus group-II; grouping;
XX diagnosis; amplification; primer; RT-PCR;
XX reverse transcription polymerase chain reaction.

OS Hepatitis C virus.

XX EF586065-A2.

XX 09-MAR-1994.

XX 16-JUL-1993; 93EP-00305591.

XX 16-JUL-1992; 92JP-00212061.

XX 30-OCT-1992; 92JP-00316634.

XX 30-OCT-1992; 92JP-00316635.

XX 30-APR-1993; 93JP-00104754.

XX (TOFU) TONEN CORP.

XX Hasegawa A, Maki N, Yagi S, Kashiwakuma T, Yamaguchi K;

XX Ikeguchi N, Kobayashi T, Sanoo C;

XX WPI; 1994-076364/10.

XX N-PSDB; AAQ58477.

XX New antigenic peptide(s) specific for hepatitis C antibodies - allowing
XX differentiation between viral gps. I and II, useful for diagnosis and
XX grouping.

XX Disclosure; Page 23; 35pp; English.

XX DNA of sequence AAQ58477 encodes antigenic peptide C14-2-2 (AAR49654),
XX used in the identification of HCV group-II strains. The DNA was obtained
XX by RT-PCR cloning of HCV genes using the primers given in AAQ58484-85.

XX (Updated on 25-MAR-2003 to correct FN field.) (Updated on 25-MAR-2003 to
XX correct FI field.)

XX Sequence 87 AA;

Query Match 5.9%; Score 7; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 23 PDKEVLY 29
|||||

RESULT 763

AAW44816
ID AAW44816 standard; protein; 88 AA.

XX AAW44816;

XX
DT 21-JUL-1998 (first entry)

XX

DE Hepatitis C virus HC88 antigenic polypeptide.
 XX Antigen; HCV; detection; antibody; infection; genetic engineering;
 KW immunology.
 XX
 OS Hepatitis C virus.
 XX
 PN RU2084527-Cl.
 XX
 XX
 PD 20-JUL-1997.
 XX
 XX
 PF 03-JUN-1993; 93RU-00029437.
 XX
 PR 03-JUN-1993; 93RU-00029437.
 XX
 PA (BIOS=) BIOSERVIS BIOTECHN CO.
 XX
 PI Kraanykh VN, Loparev VN, Blinov VM;
 XX
 DR WPI: 1998-119091/11.
 DR N-PSDB; AAV19250.
 XX
 PT Composition for detecting hepatitis C virus antibodies - comprises
 PT polypeptide(s) adsorbed on solid support, useful in, e.g. genetic
 PT engineering.
 XX
 PS Claim 1; Col 15-16; 12pp; Russian.
 XX
 CC This sequence represents an antigenic polypeptide designated HC88 derived
 CC from Hepatitis C virus (HCV). The protein may be used in compositions,
 CC along with the HCV antigens HC117 (AAW44814) and HC82 (AAW44815), for
 CC detecting antibodies to HCV, useful in, e.g. establishing the stage of
 CC the infection and methods for treatment of hepatitis C. The compositions
 CC may also be used in genetic engineering and immunology for, e.g.
 CC preparation of recombinant proteins and bacterial cells useful for
 CC production of therapeutic and diagnostic agents against HCV infection
 XX
 XX Sequence 88 AA;
 SQ
 Query Match 5.9%; Score 7; DB 2; Length 88;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 32 GCWIVG 38
 DB 6 GCWIVG 12
 DE
 RESULT 764
 ADZ14570
 ID ADZ14570 standard; protein; 93 AA.
 XX
 AC ADZ14570;
 XX
 DT 16-JUN-2005 (first entry)
 DE
 DE Human tumor associated antigenic protein (LOC285916) Seq 86.
 KW tumor-associated antigen; antisense therapy; RNA interference; diagnosis;
 KW cytostatic; cancer; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO2005030250-A2.
 XX
 PD 07-APR-2005.
 XX
 XX 23-SEP-2004; 2004WO-EP010697.
 PF
 XX 26-SEP-2003; 2003DE-01044799.
 PR
 XX (GANY-) GANYMED PHARM AG.
 PA
 PA

PI Tuercei O, Sahin U, Helftenbein G, Schlueter V;
 XX
 DR WPI: 2005-285105/29.
 DR N-PSDB; ADZ14569.
 XX
 PT Compositions for treating and diagnosing cancer, contain agents that
 PT inhibit activity or expression of specific tumor-associated antigens, or
 PT bind to these antigens or nucleic acid encoding them.
 XX
 PS Claim 32; SEQ ID NO 86; 388pp; German.
 XX
 CC This invention relates to a novel pharmaceutical composition which
 CC comprises an agent that inhibits the activity or expression of a specific
 CC tumor-associated antigen (TAG). Specifically, it relates to tumor-
 CC associated antigens that are encoded by one of the following 75 nucleic
 CC acids sequences, fragments or derivatives thereof as given in the
 CC specification. The present invention describes antisense nucleic acids
 CC that hybridize to these TAG polynucleotides that may be used for
 CC antisense therapy and RNA interference, as well as methods for diagnosing
 CC a disease associated with (abnormal) expression of TAG. Accordingly, it
 CC further relates to methods for determining regression, progression and
 CC onset of a disease by administering an antibody, optionally linked to a
 CC therapeutic or diagnostic agent, that binds to TAG. As such, cytostatic
 CC compositions derived thereof are used for treating a wide range of
 CC cancers and their metastases, where the agents that bind specifically to
 CC TAG, and the nucleic acids that encode them, are useful for diagnosis and
 CC monitoring. This polypeptide sequence is a human tumor associated
 CC antigenic protein of the invention.
 XX
 XX Sequence 93 AA;
 SQ
 Query Match 5.9%; Score 7; DB 9; Length 93;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLALAA 26
 DB 29 VLALAA 35
 DE
 RESULT 765
 ADY22233
 ID ADY22233 standard; protein; 107 AA.
 XX
 AC ADY22233;
 XX
 DT 21-APR-2005 (first entry)
 DE
 DE Plant full length insert polypeptide seqid 70017.
 XX
 KW plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.
 XX
 OS Unidentified.
 XX
 PN US2004034888-A1.
 XX
 PD 19-FEB-2004.
 XX
 XX 28-APR-2003; 2003US-00425114.
 PF
 XX 06-MAY-1999; 99US-00304517.
 PR
 XX 05-NOV-2001; 2001US-00985678.
 PR
 XX (LIUJ/) LIU J.
 PA (ZHOU/) ZHOU Y.
 PA (KOVA/) KOVALIC D K.

XX Claim 26; SEQ ID NO 4; 122pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of Tsg101
 CC associated ligase (Tal) enzymes. The DNA and protein sequences of the
 CC invention are useful for treating HIV infection and/or a
 CC hyperproliferative disease associated with dysregulated activity of
 CC Tsg101. The present amino acid sequence represents a murine Tal enzyme of
 CC the invention.
 XX
 XX Sequence 116 AA;

Query Match 5.9%; Score 7; DB 9; Length 116;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQRA 93
 |||||
 Db 103 LGLLQRA 109

RESULT 768
 AAR44407
 ID AAR44407 standard; protein; 125 AA.

XX AAR44407;
 XX
 XX 03-JUN-1994 (first entry)
 XX
 XX NANBHV derived polypeptide.

XX NANBHV; non-A non-B hepatitis virus; prophylaxis; liver; serum;
 XX chimpanzee; clone; kit.

XX Hepatitis virus.

XX JP05284969-A.

XX 02-NOV-1993.

XX 09-APR-1992; 92JP-00088840.

XX 09-APR-1992; 92JP-00088840.

XX (DAUC) DAIICHI KAGAKU YAKUHIIN KK.
 XX (DAUC) DAIICHI PHARM CO LTD.

XX WPI; 1993-382212/48.

XX Hepatitis virus gene for corresp. polypeptide - used in treatment and
 PT prophylaxis of non-A, non-B hepatitis, for encoding specified base
 PT aminoacid sequence.

XX Claim 4; Page 8; 11pp; Japanese.

XX The DNA sequences (AAQ50623-28) are obtained by extracting RNA from liver
 CC or serum of a patient or chimpanzee infected with NANBHV, synthesising
 CC cDNA and detecting the gene fragment which is negative to anti-HCV
 CC antibody and cloning the fragment. The derived proteins (AAR44404-08) can
 CC be used to detect NANBHV. The DNA and protein are useful in the treatment
 CC or prophylaxis of non-A, non-B hepatitis. The primers (AAQ50629-30) are
 CC used in the amplification process

XX Sequence 125 AA;

Query Match 5.9%; Score 7; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
 |||||
 Db 58 PDKEVLY 64

RESULT 769
 AAR44406
 ID AAR44406 standard; protein; 125 AA.
 XX
 AC AAR44406;

XX 03-JUN-1994 (first entry)
 XX
 XX NANBHV derived polypeptide.

XX NANBHV; non-A non-B hepatitis virus; prophylaxis; liver; serum;
 KW chimpanzee; clone; kit.

XX Hepatitis virus.

XX JP05284969-A.

XX 02-NOV-1993.

XX 09-APR-1992; 92JP-00088840.

XX 09-APR-1992; 92JP-00088840.

XX (DAUC) DAIICHI KAGAKU YAKUHIIN KK.
 XX (DAUC) DAIICHI PHARM CO LTD.

XX WPI; 1993-382212/48.

XX Hepatitis virus gene for corresp. polypeptide - used in treatment and
 PT prophylaxis of non-A, non-B hepatitis, for encoding specified base
 PT aminoacid sequence.

XX Claim 3; Page 7; 11pp; Japanese.

XX The DNA sequences (AAQ50623-28) are obtained by extracting RNA from liver
 CC or serum of a patient or chimpanzee infected with NANBHV, synthesising
 CC cDNA and detecting the gene fragment which is negative to anti-HCV
 CC antibody and cloning the fragment. The derived proteins (AAR44404-08) can
 CC be used to detect NANBHV. The DNA and protein are useful in the treatment
 CC or prophylaxis of non-A, non-B hepatitis. The primers (AAQ50629-30) are
 CC used in the amplification process

XX Sequence 125 AA;

Query Match 5.9%; Score 7; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
 |||||
 Db 58 PDKEVLY 64

RESULT 770
 AAR44408
 ID AAR44408 standard; protein; 125 AA.

XX AAR44408;

XX 03-JUN-1994 (first entry)

XX NANBHV derived polypeptide.

XX NANBHV; non-A non-B hepatitis virus; prophylaxis; liver; serum;
 KW chimpanzee; clone; kit.

XX Hepatitis virus.

XX JP05284969-A.

XX 02-NOV-1993.

```
PP 09-APR-1992; 92JP-00088840.
XX
PR 09-APR-1992; 92JP-00088840.
XX
PA (DAUC ) DAIICHI KAGAKU YAKUHIIN KK.
PA (DAUC ) DAIICHI PHARM CO LTD.
XX
DR WPI; 1993-382212/48.
XX
XX Hepatitis virus gene for corresp. polypeptide - used in treatment and
PT prophylaxis of non-A, non-B-hepatitis, for encoding specified base
PT aminoacid sequence.
XX
PS Claim 5; Page 8; 11pp; Japanese.
XX
XX The DNA sequences (AAQ50623-28) are obtained by extracting RNA from liver
CC or serum of a patient or chimpanzee infected with NANBH, synthesising
CC cDNA and detecting the gene fragment which is negative to anti-HCV
CC antibody and cloning the gene fragment. The derived proteins (AAR44404-08) can
CC be used to detect NANBH. The DNA and protein are useful in the treatment
CC or prophylaxis of non-A, non-B hepatitis. The primers (AAQ50629-30) are
CC used in the amplification process
XX
SQ Sequence 125 AA;

Query Match 5.9%; Score 7; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 58 PDKEVLY 64

RESULT 771
AAR44404
ID AAR44404 standard; protein; 125 AA.
XX
AC AAR44404;
XX
DT 03-JUN-1994 (first entry)
XX
DE NANBH derived polypeptide.
XX
XX NANBH; non-A non-B hepatitis virus; prophylaxis; liver; serum;
KW chimpanzee; clone; kit.
XX
OS Hepatitis virus.
XX
PN JP05284969-A.
XX
PD 02-NOV-1993.
XX
PF 09-APR-1992; 92JP-00088840.
XX
PR 09-APR-1992; 92JP-00088840.
XX
PA (DAUC ) DAIICHI KAGAKU YAKUHIIN KK.
PA (DAUC ) DAIICHI PHARM CO LTD.
XX
XX WPI; 1993-382212/48.
XX
XX Hepatitis virus gene for corresp. polypeptide - used in treatment and
PT prophylaxis of non-A, non-B-hepatitis, for encoding specified base
PT aminoacid sequence.
XX
PS Claim 2; Page 7; 11pp; Japanese.
XX
XX The DNA sequences (AAQ50623-28) are obtained by extracting RNA from liver
CC or serum of a patient or chimpanzee infected with NANBH, synthesising
CC cDNA and detecting the gene fragment which is negative to anti-HCV
CC antibody and cloning the gene fragment. The derived proteins (AAR44404-08) can
CC be used to detect NANBH. The DNA and protein are useful in the treatment
CC or prophylaxis of non-A, non-B hepatitis. The primers (AAQ50629-30) are
CC used in the amplification process
XX
SQ Sequence 125 AA;

Query Match 5.9%; Score 7; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 58 PDKEVLY 64

RESULT 772
AAR44405
ID AAR44405 standard; protein; 125 AA.
XX
AC AAR44405;
XX
DT 03-JUN-1994 (first entry)
XX
DE NANBH derived polypeptide.
XX
XX NANBH; non-A non-B hepatitis virus; prophylaxis; liver; serum;
KW chimpanzee; clone; kit.
XX
OS Hepatitis virus.
XX
PN JP05284969-A.
XX
PD 02-NOV-1993.
XX
PF 09-APR-1992; 92JP-00088840.
XX
PR 09-APR-1992; 92JP-00088840.
XX
PA (DAUC ) DAIICHI KAGAKU YAKUHIIN KK.
PA (DAUC ) DAIICHI PHARM CO LTD.
XX
XX WPI; 1993-382212/48.
XX
XX Hepatitis virus gene for corresp. polypeptide - used in treatment and
PT prophylaxis of non-A, non-B-hepatitis, for encoding specified base
PT aminoacid sequence.
XX
PS Claim 2; Page 7; 11pp; Japanese.
XX
XX The DNA sequences (AAQ50623-28) are obtained by extracting RNA from liver
CC or serum of a patient or chimpanzee infected with NANBH, synthesising
CC cDNA and detecting the gene fragment which is negative to anti-HCV
CC antibody and cloning the gene fragment. The derived proteins (AAR44404-08) can
CC be used to detect NANBH. The DNA and protein are useful in the treatment
CC or prophylaxis of non-A, non-B hepatitis. The primers (AAQ50629-30) are
CC used in the amplification process
XX
SQ Sequence 125 AA;

Query Match 5.9%; Score 7; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 58 PDKEVLY 64

RESULT 773
AAR63300
ID AAR63300 standard; protein; 128 AA.
XX
AC AAR63300;
XX
```

DT 25-MAR-2003 (revised)
 DT 07-AUG-1995 (first entry)
 XX Polypeptide encoded by hepatitis C virus NS3 sequence.
 XX Hepatitis C virus; HCV; primer; probe; detection; diagnosis;
 KW classification; immunisation; prophylaxis; serotyping.
 XX
 XX Hepatitis C virus type 5a.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 13 /note= "Unspecified amino acid."
 FT Misc-difference 26 /note= "Unspecified amino acid."
 FT Misc-difference 27 /note= "Unspecified amino acid."
 FT Misc-difference 51 /note= "Unspecified amino acid."
 FT Misc-difference 76 /note= "Unspecified amino acid."
 FT Misc-difference 117 /note= "Unspecified amino acid."
 FT Misc-difference 117 /note= "Unspecified amino acid."
 XX
 PN W09425601-A2.
 XX
 XX 10-NOV-1994.
 XX
 XX 27-APR-1994; 94WO-EP001323.
 XX
 XX 27-APR-1993; 93EP-00401099.
 PR
 PR 05-AUG-1993; 93EP-00402019.
 XX
 XX (INNO-) INNOGENETICS NV SA.
 XX
 XX Maertens G, Stuyver L;
 XX
 XX WPI; 1994-358277/44.
 DR N-PSDB; AAQ78052.
 XX
 XX New polynucleotide sequences from hepatitis C virus - and related
 PT vectors, polypeptide(s) and antibodies, useful for immunisation,
 PT treatment, diagnosis and typing of HCV isolates.
 XX
 XX Disclosure; Page 149; 404pp; English.
 XX
 XX Compositions comprising at least 5, and pref. 8 or more contiguous
 CC nucleotides selected from an HCV type 3 genomic sequence, more
 CC particularly (i) the region spanning positions 417-957 of the Core/E1
 CC region of HCV subtype 3a; (ii) the region spanning positions 4664-4730 of
 CC the NS3 region of HCV type 3; (iii) the region spanning positions 4892-
 CC 5292 of the NS3/4 region of HCV type 3; (iv) the region spanning
 CC positions 8023-8235 of the NS5 region of the BR36 subgroup of HCV subtype
 CC 3a; or (v) an HCV subtype 3c genomic sequence, or, from a subtype 2d
 CC genomic sequence, a type 4 genomic sequence; or the coding region of
 CC subtype 5a, may be used as primers to amplify nucleic acid from an
 CC isolate belonging to a specific genotype, or as a probe for specific
 CC detection/classification of nucleic acid. Polypeptides encoded by the
 CC nucleotides in such compositions may be used for immunisation against
 CC HCV, for the detection of antibodies directed against HCV and for
 CC serotyping. This sequence corresponds to the NS3 region of HCV subtype 5a
 CC and is taken from a clone designated PC-1-37. (Updated on 25-MAR-2003 to
 CC correct PN field.)
 XX
 XX Sequence 128 AA;
 SQ
 Query Match 5.9%; Score 7; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMEECS 66
 DB 61 DEMEECS 67
 DT 10-AUG-2000 (first entry)

RESULT 774
 ABO76790
 ID ABO76790 standard; protein; 130 AA.
 XX
 XX ABO76790;
 AC
 XX 29-JUL-2004 (first entry)
 DT
 XX Pseudomonas aeruginosa polypeptide #8965.
 DE
 XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 KW
 XX Pseudomonas aeruginosa.
 OS
 XX US6551795-B1.
 PN
 XX 22-APR-2003.
 PD
 XX 18-FEB-1999; 99US-00252991.
 PF
 XX 18-FEB-1998; 98US-0074788P.
 PR
 PR 27-JUL-1998; 98US-0094190P.
 XX
 XX (GENO-) GENOME THERAPEUTICS CORP.
 PA
 XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 PI
 XX WPI; 2003-615309/58.
 DR
 DR N-PSDB; ABD10361.
 XX
 XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 XX Disclosure; SEQ ID NO 25536; 455pp; English.
 PS
 XX The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 XX Sequence 130 AA;
 SQ
 Query Match 5.9%; Score 7; DB 7; Length 130;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 LEVTTST 13
 DB 49 LEVTTST 55

RESULT 775
 AAY92483
 ID AAY92483 standard; protein; 131 AA.
 XX
 AC AAY92483;
 XX
 DT 10-AUG-2000 (first entry)

XX TYLCV C3 mutant, mc3#47.
XX C3; AL3; mutant; transgenic plant; resistance; geminivirus; infection.
XX Tomato yellow leaf curl virus.
XX WO200020614-A1.
XX 13-APR-2000.
XX 02-SEP-1999; 99WO-US020164.
XX 01-OCT-1998; 98US-00164615.
XX (UTNC-) UNIV NORTH CAROLINA STATE.
XX Bowdoin LH, Settlage S;
XX WPI; 2000-303792/26.
XX Plant comprises transformed plant cells containing heterologous nucleic acid construct for increasing resistance to geminivirus.
XX Example 1; Page 65; 82pp; English.
XX AAY92467-97 are C3 mutants from Tomato yellow leaf curl virus. The designations C3 and AL3 are interchangeable - viral genes are named as to whether they are specified by the virion (V) or complementary (C) sense DNA strands, or alternatively with respect to the left (L) or right (R) of the 5' intergenic region. Plants, containing transformed plant cells containing a heterologous nucleic acid construct comprising a mutant AL3/C3 protein, operably associated with a promoter and a termination sequence, are claimed. Expression of the mutant AL3/C3 protein increases resistance of the plant to infection by at least one geminivirus compared to a non-transformed control. This is especially useful for increasing the resistance of cassava, potato, bean, squash, beet or tomato plants to infection by, e.g. tomato golden mosaic virus, tomato mottle virus and tomato yellow leaf curl virus
XX Sequence 131 AA;
Query Match 5.9%; Score 7; DB 3; Length 131;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 GVLAALA 25
Db 86 GVLAALA 92
|||
RESULT 776
ADE08302
ID ADE08302 standard; protein; 145 AA.
AC ADE08302;
XX 29-JAN-2004 (first entry)
XX Novel protein (useful for identifying genetic disorders) #457.
XX novel gene; novel protein; tissue marker; molecular weight marker;
XX chromosome marker; genetic disorder.
XX Unidentified.
XX WO20003054152-A2.
XX 03-JUL-2003.
XX 10-DEC-2002; 2002WO-US039555.
XX 10-DEC-2001; 2001US-0339739P.

PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Auendi V, Goodrich RW, Ren F, Zhang J, Zhao QX, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
XX Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
XX N-PSDB; ADE07391.
XX New polynucleotides, useful for expressing recombinant proteins for analysis, characterization or therapeutic use, or as markers for tissues in which the corresponding protein is preferentially expressed.
XX Claim 20; SEQ ID NO 1368; 1177pp; English.
XX The invention comprises the amino acid and coding sequences of novel proteins. The DNA and protein sequences of the invention are useful as: markers for tissues in which the corresponding protein is preferentially expressed; as molecular weight markers on gels; as chromosome markers or tags; to identify chromosomes or to map related gene positions; and to compare with endogenous DNA sequences in patients to identify potential genetic disorders. The present amino acid sequence represents a protein of the invention.
XX Sequence 145 AA;
Query Match 5.9%; Score 7; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 32 GCVVIVG 38
Db 54 GCVVIVG 60
|||
RESULT 777
AAB76733
ID AAB76733 standard; protein; 146 AA.
XX AAB76733;
XX 11-APR-2001 (first entry)
XX Corynebacterium glutamicum MCT protein SEQ ID NO:448.
XX Corynebacterium glutamicum; brevibacterium lactofermentum; MCT; membrane construction and membrane transport protein; petroleum spill; hydrocarbon degradation; gram positive aerobic bacterium; marker; identification; microorganism; fine chemical production; transformation; genome mapping; genetic engineering.
XX Corynebacterium glutamicum.
XX WO200100805-A2.
XX 04-JAN-2001.
XX 23-JUN-2000; 2000WO-IB000926.
XX 25-JUN-1999; 99US-0141031P.
XX 08-JUL-1999; 99DE-01031454.
XX 08-JUL-1999; 99DE-01031478.
XX 08-JUL-1999; 99DE-01031563.
XX 09-JUL-1999; 99DE-01032122.
XX 09-JUL-1999; 99DE-01032124.

```

PR 09-JUL-1999; 99DE-01032125.
PR 09-JUL-1999; 99DE-01032128.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032182.
PR 09-JUL-1999; 99DE-01032190.
PR 09-JUL-1999; 99DE-01032191.
PR 09-JUL-1999; 99DE-01032209.
PR 09-JUL-1999; 99DE-01032212.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032228.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032927.
PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040830.
PR 27-AUG-1999; 99DE-01040831.
PR 27-AUG-1999; 99DE-01040832.
PR 27-AUG-1999; 99DE-01040833.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041395.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042078.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042088.
XX
PA (BADI ) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR WPI; 2001-071486/08.
XX
DR N-PSDB; AAF67966.
XX
PT Corynebacterium glutamicum nucleic acids encoding membrane construction
PT and membrane transport proteins or their portions, useful for typing or
PT identifying C. glutamicum or related bacteria, and as markers for
PT transformation.
XX
PS Claim 20; Page 794; 1119pp; English.
XX
XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
CC construction and membrane transport (MCT) proteins given in AAB76510 to
CC AAB76847. The MCT nucleic acids and proteins are useful in the
CC identification of microorganisms which can be used to produce fine
CC chemicals, for modulating fine chemical production in C. glutamicum or
CC related bacteria (e.g. Brevibacterium lactofermentum), the typing or
CC identification of C. glutamicum or related bacteria, as reference points
CC for mapping C. glutamicum genome, and as markers for transformation.
CC AAF68082 and AAF68082 represent sequencing primers which are used in an
CC example from the present invention
XX
SQ Sequence 146 AA;

Query Match 5.9%; Score 7; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRAT 94
DB 58 GLLQRAT 64
|||||

RESULT 778
ADM26550
ID ADM26550 standard; protein; 147 AA.
XX
AC ADM26550;
XX
DT 20-MAY-2004 (first entry)

Query Match 5.9%; Score 7; DB 7; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 7 VLAALAA 13
|||||

RESULT 779
ADE08303
ID ADE08303 standard; protein; 153 AA.
XX
AC ADE08303;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel protein (useful for identifying genetic disorders) #458.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder.
XX
OS Unidentified.
XX
PN WO2003054152-A2.
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.

```

```

XX
DE Hyperthermophile Methanopyrus kandleri protein #1156.
XX
KW hyperthermophile; protein stability enhancement;
KW Protein activity enhancement.
XX
OS Methanopyrus kandleri.
XX
PN WO2003076575-A2.
XX
PD 18-SEP-2003.
XX
PF 04-MAR-2003; 2003WO-US006664.
XX
PR 04-MAR-2002; 2002US-0361742P.
PR 14-MAY-2002; 2002US-0380423P.
PR 16-SEP-2002; 2002US-0410974P.
XX
XX (FIDE-) FIDELITY SYSTEMS INC.
PA (MALY/) MALYKH A.
XX
PI Slesarev AI, Pavlov A, Pavlova N, Kozyavkin S;
XX
DR WPI; 2003-748383/70.
DR N-PSDB; ADM27081.
XX
PT New isolated nucleic acids encoding any of about 1700 Methanopyrus
PT kandleri proteins, and the encoded proteins, useful as a medicaments or
PT as diagnostic agents.
XX
PS Claim 31; SEQ ID NO 1156; 1023pp; English.
XX
CC The invention comprises the amino acid sequence of proteins from the
CC hyperthermophile Methanopyrus kandleri, the invention also comprises the
CC complete genome from Methanopyrus kandleri. The Methanopyrus kandleri
CC proteins of the invention are useful for enhancing the stability and/or
CC activity of other proteins. The Methanopyrus kandleri genome is useful in
CC a variety of diagnostic and analytical methods. The present amino acid
CC sequence represents a Methanopyrus kandleri protein of the invention.
XX
SQ Sequence 147 AA;

Query Match 5.9%; Score 7; DB 7; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 7 VLAALAA 13
|||||

RESULT 779
ADE08303
ID ADE08303 standard; protein; 153 AA.
XX
AC ADE08303;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel protein (useful for identifying genetic disorders) #458.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder.
XX
OS Unidentified.
XX
PN WO2003054152-A2.
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.

```

PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365031P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
DR WPI; 2003-569235/53.
DR N-PSDB; ADE07392.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
XX Claim 20; SEQ ID NO 1369; 1177pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence represents a protein
CC of the invention.
XX
XX
SQ Sequence 153 AA;
Query Match 5.9%; Score 7; DB 7; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 32 GCVVIVG 38
Db 54 GCVVIVG 60
|||||
RESULT 780
ADX78761
ID ADX78761 standard; protein; 161 AA.
AC
AC ADX78761;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polypeptide seqid 48127.
DE
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
XX Unidentified.
OS
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
PD
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX

PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABAY/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 48127; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 161 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 LLGGVLA 22
Db 87 LLGGVLA 93
|||||
RESULT 781
ABO81086
ID ABO81086 standard; protein; 163 AA.
XX
XX ABO81086;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #13261.
DE
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KW
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX 27-JUL-1998; 98US-0094190P.
PR
XX (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 DR WPI; 2003-615309/58.
 XX N-PSDB; ABD14657.
 XX Novel isolated nucleic acid encoding *Pseudomonas aeruginosa* polypeptide, of
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX Disclosure; SEQ ID NO 29832; 455pp; English.
 XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics, and
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-*P. aeruginosa* drugs, as templates for recombinant
 CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
 CC infection, and in detection of *P. aeruginosa* sequences or other sequences
 CC of *Pseudomonas* species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent *P. aeruginosa* polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 163 AA;

Query Match 5.9%; Score 7; DB 7; Length 163;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LGGVLA 22
 |||||
 Db 154 LGGVLA 160

RESULT 782

ADX71619
 ID ADX71619 standard; protein; 163 AA.

AC ADX71619;

DT 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 40985.

XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

XX Unidentified.

OS US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

PR 05-NOV-2001; 2001US-00985678.

XX (LIU/J) LIU J.

PA (ZHOU/J) ZHOU Y.

PA (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.

PA (TABA/) TABASKA J E.
 XX (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JB, Cao Y;

XX WPI; 2004-180133/17.

DR New recombinant DNA construct, useful for improving plant tolerance to
 XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

XX Claim 1; SEQ ID NO 40985; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.seqdata.uspto.gov/sequence.html?DocId:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX SQ Sequence 163 AA;

Query Match 5.9%; Score 7; DB 8; Length 163;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LGGVLA 22
 |||||
 Db 87 LGGVLA 93

RESULT 783

ADY06215
 ID ADY06215 standard; protein; 171 AA.

XX AC ADY06215;

XX 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 62030.

XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

PR 05-NOV-2001; 2001US-00985678.

XX PA (LIUJ/) LIU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABAS/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX XX WPI; 2004-180133/17.
XX XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX XX Claim 1; SEQ ID NO 62030; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
XX invention.
XX SQ Sequence 171 AA;
Query Match 5.9%; Score 7; DB 8; Length 171;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 16 LGGVLA 22
Db 108 LGGVLA 114
|||||
RESULT 784
AAG75519
ID AAG75519 standard; protein; 172 AA.
XX AC AAG75519;
XX DT 03-SEP-2001 (first entry)
XX DE Human colon cancer antigen protein SEQ ID NO:6283.
XX KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma.
XX OS Homo sapiens.
XX PN WO200122920-A2.
XX PD 05-APR-2001.
XX PF 28-SEP-2000; 2000WO-US026524.
XX PR 29-SEP-1999; 99US-0157137P.
XX PR 03-NOV-1999; 99US-0163280P.
XX PA (LIUJ-) HUMAN GENOME SCI INC.
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX XX WPI; 2001-235357/24.
XX DR N-PSDB; AAH34924.
XX XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX XX Claim 11; Page 7736-7737; 9803pp; English.
XX CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX SQ Sequence 172 AA;
Query Match 5.9%; Score 7; DB 4; Length 172;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 18 GGVLAAAL 24
Db 125 GGVLAAAL 131
|||||
RESULT 785
AAM06677
ID AAM06677 standard; protein; 173 AA.
XX AC AAM06677;
XX DT 05-OCT-2001 (first entry)
XX DE Human foetal protein, SEQ ID NO: 408.
XX KW Human; foetal protein; cytostatic; immunosuppressive; immunostimulant;
KW neutropic; neuroprotective; thrombolytic; osteopathic; antiinflammatory;
KW gene therapy; antitense therapy; cancer; immune disorder;
KW growth disorder; osteoporosis; thrombolytic disorder;
KW nervous system disorder; inflammation.
XX OS Homo sapiens.
XX PN WO200155339-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US002723.
XX PR 25-JAN-2000; 2000US-00491404.
XX PR 15-SEP-2000; 2000US-00663870.
XX PR 06-NOV-2000; 2000US-00707351.
XX PA (HYSE-) HYSEQ INC.
XX PI Yeung G, Ford JE, Boyle BJ, Arterburn MC, Drmanac RA, Tang YT;

PI Liu C, Asundi V, Zhou P, Werhman T;
XX WPI; 2001-465571/50.
DR N-PSDB; AAH94352.
XX
PT Novel fetal proteins useful for the treatment and diagnosis of diseases
PT associated with dysfunction of the protein e.g. cancers, immune
PT disorders, growth disorders, thrombolytic disorders, nervous system
PT disorders and inflammation.
XX
PS Claim 10; Page 326; 715pp; English.
XX
CC The invention relates to novel foetal polypeptides encoded by
CC polynucleotides comprising one of 477 sequences fully defined in the
CC specification. The foetal polynucleotides and polypeptides are useful in
CC the treatment and diagnosis of diseases such as cancers, immune
CC disorders, growth disorders (e.g. osteoporosis), thrombolytic disorders,
CC nervous system disorders and inflammation. The present sequence is a
CC polypeptide encoded by a cDNA assembled using an expressed sequence tag
CC (EST) found to be expressed in human foetal tissue cDNA libraries
XX
SQ Sequence 173 AA;
Query Match 5.9%; Score 7; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 109 QKLEAFW 115
DB 118 QKLEAFW 124
|||||
RESULT 786
ADX94276
ID ADX94276 standard; protein; 177 AA.
XX
AC ADX94276;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 56940.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
FN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX

PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 56940; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.secdatasupto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 177 AA;
Query Match 5.9%; Score 7; DB 8; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAL 24
DB 65 GGVLAAL 71
|||||
RESULT 787
ADY04940
ID ADY04940 standard; protein; 177 AA.
XX
AC ADY04940;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polypeptide seqid 60755.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
OS Unidentified.
XX
FN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX

XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX DR WPI; 2004-180133/17.
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX PT pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PS Claim 1; SEQ ID NO 60755; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX CC available in electronic form from the US patent office at
XX CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX CC of the invention are also useful in physical arrays of molecules and as
XX CC plant breeding markers. The recombinant DNA construct is useful for
XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX CC plant cells by modification of the cell cycle pathway, for conferring
XX CC increased resistance to plant disease, for producing galactomannan,
XX CC lignin or plant growth regulators, for increasing the rate of homologous
XX CC recombination in plants, for improving yield by modification of
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX CC or by providing improved plant growth and development under at least one
XX CC stress condition or for modifying seed oil or protein yield and/or
XX CC content. This is the amino acid sequence of a plant full length insert
XX CC polypeptide that can be used in the recombinant DNA construct of the
XX CC invention.
XX SQ Sequence 177 AA;
Query Match 5.9%; Score 7; DB 8; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 65 GGVLAAL 71
RESULT 788
ADY22585
ID ADY22585 standard; protein; 177 AA.
AC ADY22585;
XX 21-APR-2005 (first entry)
XX Plant full length insert polypeptide seqid 70369.
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX OS Unidentified.
XX US2004034888-A1.
XX 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX PT pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PS Claim 1; SEQ ID NO 70369; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX CC available in electronic form from the US patent office at
XX CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX CC of the invention are also useful in physical arrays of molecules and as
XX CC plant breeding markers. The recombinant DNA construct is useful for
XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX CC plant cells by modification of the cell cycle pathway, for conferring
XX CC increased resistance to plant disease, for producing galactomannan,
XX CC lignin or plant growth regulators, for increasing the rate of homologous
XX CC recombination in plants, for improving yield by modification of
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX CC or by providing improved plant growth and development under at least one
XX CC stress condition or for modifying seed oil or protein yield and/or
XX CC content. This is the amino acid sequence of a plant full length insert
XX CC polypeptide that can be used in the recombinant DNA construct of the
XX CC invention.
XX SQ Sequence 177 AA;
Query Match 5.9%; Score 7; DB 8; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 65 GGVLAAL 71
RESULT 789
RAY41260
ID AAY41260 standard; protein; 179 AA.
AC AAY41260;
XX 31-JAN-2000 (first entry)
XX Amino acid sequence of short murine FAIM.
XX Fas Apoptosis Inhibitory Molecule; apoptosis; Bal-17 B lymphoma cell;
XX murine; faim; FAIM.
XX OS Mus sp.
XX WO9953743-A2.
XX 28-OCT-1999.
XX 20-APR-1999; 99WO-US008658.
XX 21-APR-1998; 98US-0082503P.
XX 15-MAR-1999; 99US-0124805P.
XX (UYBO-) UNIV BOSTON.
PA (UYBO-) UNIV BOSTON.

PI Rothstein TL, Schneider TJ, Donohoe TJ;
XX WPI; 1999-633871/54.
DR N-PSDB; AAZ28043.
XX
XX
XX Fas Apoptosis Inhibitory Molecule (FAIM) oligonucleotides and proteins.
XX
XX Disclosure; Fig 2J; 43pp; English.
XX
XX The invention relates to nucleic acid sequences from the Fas Apoptosis
CC Inhibitory Molecule (FAIM) gene, and the encoded proteins. The proteins
CC prevent apoptosis in Bal-17 B lymphoma cells and various murine cells.
CC Host cells expressing one of the FAIM nucleic acid sequence are used for
CC screening for compounds that alter FAIM activity. These compounds are
CC useful for modulating apoptosis. The antibodies (and any fusion proteins
CC produced) are useful for identifying novel FAIM pathway constituents and
CC FAIM homologs. The antibodies are also useful for determining the level
CC of FAIM expression in various cell types. FAIM inhibits apoptosis when
CC the gene is expressed ectopically. The present sequence represents the
CC amino acid sequence of short murine FAIM
XX
XX Sequence 179 AA;
SQ
Query Match 5.9%; Score 7; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 TTSTWVL 16
Db 96 TTSTWVL 102
|||||
RESULT 790
ADX78726
ID ADX78726 standard; protein; 179 AA.
XX
XX ADX78726;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polypeptide seqid 48092.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
XX Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX

PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 48092; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.secdatasupto.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
XX invention.
XX
XX Sequence 179 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 LLGGVLA 22
Db 108 LLGGVLA 114
|||||
RESULT 791
ADX78414
ID ADX78414 standard; protein; 179 AA.
XX
XX ADX78414;
XX
XX 21-APR-2005 (first entry)
XX
XX Plant full length insert polypeptide seqid 47780.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX
XX Unidentified.
XX
XX US2004034888-A1.
XX
XX 19-FEB-2004.
XX
XX 28-APR-2003; 2003US-00425114.
XX
XX 06-MAY-1999; 99US-00304517.
XX
XX 05-NOV-2001; 2001US-00985678.
XX
XX (LIUJ/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX

XX PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX DR WPI; 2004-180133/17.
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX PT pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PS Claim 1; SEQ ID NO 47780; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX CC available in electronic form from the US patent office at
XX CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX CC of the invention are also useful in physical arrays of molecules and as
XX CC plant breeding markers. The recombinant DNA construct is useful for
XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX CC plant cells by modification of the cell cycle pathway, for conferring
XX CC increased resistance to plant disease, for producing galactomannan,
XX CC lignin or plant growth regulators, for increasing the rate of homologous
XX CC recombination in plants, for improving yield by modification of
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX CC or by providing improved plant growth and development under at least one
XX CC stress condition or for modifying seed oil or protein yield and/or
XX CC content. This is the amino acid sequence of a plant full length insert
XX CC polypeptide that can be used in the recombinant DNA construct of the
XX CC invention.
XX SQ Sequence 179 AA;
Query Match 5.9%; Score 7; DB 8; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 65 GGVLAAL 71
RESULT 792
ADX78722
ID ADX78722 standard; protein; 180 AA.
AC ADX78722;
XX 21-APR-2005 (first entry)
XX Plant full length insert polypeptide seqid 48088.
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX Unidentified.
XX OS US2004034888-A1.
XX PN 19-FEB-2004.
XX PD 28-APR-2003; 2003US-00425114.
XX PF 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX PX (LIUJ/) LIU J.

PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOV/) CAO Y.
XX PT Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX DR WPI; 2004-180133/17.
XX PT New recombinant DNA construct, useful for improving plant tolerance to
XX PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX PT pests, for conferring increased resistance to plant disease, or for
XX PT improving yield.
XX PS Claim 1; SEQ ID NO 48088; 15pp; English.
XX CC The invention describes a recombinant DNA construct comprising a
XX CC polynucleotide consisting of a sequence encoding an amino acid sequence
XX CC available in electronic form from the US patent office at
XX CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX CC of the invention are also useful in physical arrays of molecules and as
XX CC plant breeding markers. The recombinant DNA construct is useful for
XX CC improving plant tolerance to cold, heat, drought, herbicides, extreme
XX CC osmotic conditions, pathogens or pests, for manipulating growth rate in
XX CC plant cells by modification of the cell cycle pathway, for conferring
XX CC increased resistance to plant disease, for producing galactomannan,
XX CC lignin or plant growth regulators, for increasing the rate of homologous
XX CC recombination in plants, for improving yield by modification of
XX CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
XX CC or by providing improved plant growth and development under at least one
XX CC stress condition or for modifying seed oil or protein yield and/or
XX CC content. This is the amino acid sequence of a plant full length insert
XX CC polypeptide that can be used in the recombinant DNA construct of the
XX CC invention.
XX SQ Sequence 180 AA;
Query Match 5.9%; Score 7; DB 8; Length 180;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 74 GGVLAAL 80
RESULT 793
ADX78423
ID ADX78423 standard; protein; 180 AA.
AC ADX78423;
XX 21-APR-2005 (first entry)
XX Plant full length insert polypeptide seqid 47789.
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content.
XX Unidentified.
XX OS US2004034888-A1.
XX PN 19-FEB-2004.
XX PD 28-APR-2003; 2003US-00425114.
XX PF

XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX
XX
XX
PA (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREEN S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
XX Claim 1; SEQ ID NO 47789; 15pp; English.
XX
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 180 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 180;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 LGGVLA 22
Db 108 LGGVLA 114
RESULT 794
ABM92257
ID ABM92257 standard; protein; 183 AA.
XX
XX AC ABM92257;
XX
XX DT 02-JUN-2005 (first entry)
XX
XX DE M. xanthus protein sequence, seq id 11456.
XX
XX KW Transgenic plant; DNA replication; gene regulation; gene expression.
XX
XX OS Myxococcus xanthus.
XX
XX PN US6833447-B1.
XX
XX PD 21-DEC-2004.
XX
XX PF 10-JUL-2001; 2001US-00902540.
XX
XX PR 10-JUL-2000; 2000US-0217883P.
XX
XX PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX
XX WPI; 2005-028716/03.
XX
XX New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 11456; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16925 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX

XX (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX
XX WPI; 2005-028716/03.
XX
XX New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 11456; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16925 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
XX Sequence 183 AA;
SQ
Query Match 5.9%; Score 7; DB 9; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 81 QFKGKVL 87
Db 22 QFKGKVL 28
RESULT 795
ABM90870
ID ABM90870 standard; protein; 188 AA.
XX
XX AC ABM90870;
XX
XX DT 02-JUN-2005 (first entry)
XX
XX DE M. xanthus protein sequence, seq id 10069.
XX
XX KW Transgenic plant; DNA replication; gene regulation; gene expression.
XX
XX OS Myxococcus xanthus.
XX
XX PN US6833447-B1.
XX
XX PD 21-DEC-2004.
XX
XX PF 10-JUL-2001; 2001US-00902540.
XX
XX PR 10-JUL-2000; 2000US-0217883P.
XX
XX PA (MONS) MONSANTO TECHNOLOGY LLC.
XX
XX PI Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;
XX
XX WPI; 2005-028716/03.
XX
XX New substantially purified Myxococcus xanthus nucleic acid molecule
PT encoding a nitrite reductase, useful for determining gene expression,
PT identifying mutations in a gene of interest, and for constructing
PT mutations in a gene of interest.
XX
XX Example 2; SEQ ID NO 10069; 25pp; English.
XX
XX The invention relates to a substantially purified nucleic acid molecule
CC encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a

CC recombinant DNA construct for expression of a nitrite reductase gene in a
CC plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
CC mutations in a gene of interest, and for constructing mutations in a gene
CC of interest. Sequences given in records for SEQ IDs 9692-16825 represent
CC a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX
SQ Sequence 188 AA;

Query Match 5.9%; Score 7; DB 9; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRATQ 95
DB 95 LLQRATQ 101
|||||

RESULT 796
AAV41261
ID AAV41261 standard; protein; 201 AA.
XX
AC AAV41261;

DT 31-JAN-2000 (first entry)

DE Amino acid sequence of long murine FAIM.

KW Fas Apoptosis Inhibitory Molecule; apoptosis; Bal-17 B lymphoma cell;
KW murine; faim; FAIM.

OS Mus sp.

PN WO9953743-A2.

XX 28-OCT-1999.

XX 20-APR-1999; 99WO-US008658.

XX 21-APR-1998; 98US-0082503P.

PR 15-MAR-1999; 99US-0124805P.

XX (UYBO-) UNIV BOSTON.

XX Rothstein TL, Schneider TJ, Donohoe TJ;

XX WPI; 1999-633871/54.

XX N-PSDB; AA228044.

XX Fas Apoptosis Inhibitory Molecule (FAIM) oligonucleotides and proteins.

PS Disclosure; Fig 2L; 43pp; English.

XX The invention relates to nucleic acid sequences from the Fas Apoptosis
CC Inhibitory Molecule (FAIM) gene, and the encoded proteins. The proteins
CC prevent apoptosis in Bal-17 B lymphoma cells and various murine cells.
CC Host cells expressing one of the FAIM nucleic acid sequence are used for
CC screening for compounds that alter FAIM activity. These compounds are
CC useful for modulating apoptosis. The antibodies (and any fusion proteins
CC produced) are useful for identifying novel FAIM pathway constituents and
CC FAIM homologs. The antibodies are also useful for determining the level
CC of FAIM expression in various cell types. FAIM inhibits apoptosis when
CC the gene is expressed ectopically. The present sequence represents the
CC amino acid sequence of long murine FAIM

XX Sequence 201 AA;

Query Match 5.9%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 TTSTWVL 16
DB 118 TTSTWVL 124
|||||

RESULT 797

ADX93690
ID ADX93690 standard; protein; 202 AA.

XX
AC ADX93690;

DT 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 56354.

XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.

XX Unidentified.

XX US2004034888-A1.

PN 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

PR 05-NOV-2001; 2001US-00985678.

XX (LIUJ/) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOVA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TABR/) TABASKA J E.

XX (CAOY/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
CC cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
CC pests, for conferring increased resistance to plant disease, or for
CC improving yield.

XX Claim 1; SEQ ID NO 56354; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 202 AA;

AC	AB867469;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 29199.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
KX	
XX	Drosophila melanogaster.
OS	
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
Pf	23-MAR-2001; 2001WO-US009231.
XX	
PR	23-MAR-2000; 2000US-0191637P.
XX	
PR	11-JUL-2000; 2000US-00614150.
XX	
PA	(PEKE) PE CORP NY.
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
XX	
DR	N-PSDB; ABLI1572.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
XX	
PS	Disclosure; SEQ ID NO 29199; 2lpp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 210 AA;
Query Match	5.9%; Score 7; DB 4; Length 210;
Best Local Similarity	100.0%; Pred. No. 3.1e+02;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	19 GVLAALA 25
Db	23 GVLAALA 29
RESULT 800	
AAY75063	
ID	AAY75063 standard; protein; 217 AA.
XX	
AC	AAY75063;
XX	
DT	21-MAR-2000 (first entry)
XX	
DE	Neisseria meningitidis ORF 548 protein sequence SEQ ID NO:1600.
XX	
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine; antigenic; diagnosis; immunogenic; infection; meningitis; septicemia; antibacterial; gene therapy.
KX	
XX	Neisseria meningitidis.
OS	
XX	
PN	WO9957280-A2.
XX	

AC	AB867469;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Drosophila melanogaster polypeptide SEQ ID NO 29199.
XX	
KW	Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.
KX	
XX	Drosophila melanogaster.
OS	
XX	
PN	WO200171042-A2.
XX	
PD	27-SEP-2001.
XX	
Pf	23-MAR-2001; 2001WO-US009231.
XX	
PR	23-MAR-2000; 2000US-0191637P.
XX	
PR	11-JUL-2000; 2000US-00614150.
XX	
PA	(PEKE) PE CORP NY.
PI	Venter JC, Adams M, Li PWD, Myers EW;
XX	
DR	WPI; 2001-656860/75.
XX	
DR	N-PSDB; ABLI1572.
XX	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.
XX	
PS	Disclosure; SEQ ID NO 29199; 2lpp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 210 AA;
Query Match	5.9%; Score 7; DB 4; Length 210;
Best Local Similarity	100.0%; Pred. No. 3.1e+02;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	19 GVLAALA 25
Db	23 GVLAALA 29
RESULT 800	
AAY75063	
ID	AAY75063 standard; protein; 217 AA.
XX	
AC	AAY75063;
XX	
DT	21-MAR-2000 (first entry)
XX	
DE	Neisseria meningitidis ORF 548 protein sequence SEQ ID NO:1600.
XX	
KW	Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine; antigenic; diagnosis; immunogenic; infection; meningitis; septicemia; antibacterial; gene therapy.
KX	
XX	Neisseria meningitidis.
OS	
XX	
PN	WO9957280-A2.
XX	

Query Match	5.9%; Score 7; DB 8; Length 202;
Best Local Similarity	100.0%; Pred. No. 3e+02;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	18 GGVLAAAL 24
Db	88 GGVLAAAL 94
RESULT 798	
ADK46560	
ID	ADK46560 standard; protein; 209 AA.
XX	
AC	ADK46560;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Streptococcus pneumoniae protein, Seq ID No 3075.
XX	
KW	Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX	
OS	Streptococcus pneumoniae.
XX	
PN	US6699703-B1.
XX	
PD	02-MAR-2004.
XX	
Pf	26-MAY-2000; 2000US-00583110.
XX	
PR	02-JUL-1997; 97US-0051553P.
XX	
PR	12-MAY-1998; 98US-0085131P.
XX	
PR	30-JUN-1998; 98US-00107433.
XX	
PA	(GENO-) GENOME THERAPEUTICS CORP.
XX	
PI	Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX	
DR	WPI; 2004-212399/20.
XX	
DR	N-PSDB; ADK43899.
XX	
PT	New nucleic acid molecules and polypeptides useful for diagnosing, preventing and treating pathological conditions resulting from bacterial infection, e.g. Streptococcus pneumoniae infection, and in drug screening.
XX	
PS	Disclosure; SEQ ID NO 3075; 30lpp; English.
XX</	

XX 01-MAY-1998; 98US-0083758P.
 PR 31-JUL-1998; 98US-0094869P.
 PR 02-SEP-1998; 98US-0098994P.
 PR 02-SEP-1998; 98US-0099062P.
 PR 09-OCT-1998; 98US-0103749P.
 PR 09-OCT-1998; 98US-0103794P.
 PR 09-OCT-1998; 98US-0103796P.
 PR 25-FEB-1999; 99US-0121528P.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
 PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
 PI Tettelin H, Venter JC;
 XX WPI; 2000-062150/05.
 DR N-PSDB; AA253824.
 XX Novel Neisserial polypeptides predicted to be useful antigens for
 PT vaccines and diagnostics.
 XX Claim 2; Page 826; 1453pp; English.
 XX AA253015 to AA254536, AA254577 to AA254615, and AA274253 to AA275941
 CC represent novel *Neisseria meningitidis* and *N. gonorrhoeae* polynucleotides
 CC and polypeptides. AA254537 to AA254576 and AA254616 to AA255473 represent
 CC PCR primers used in the exemplification of the present invention. The
 CC polypeptides, the polynucleotides, antibodies and compositions of the
 CC invention can be used as vaccines, as diagnostic reagents, and as
 CC immunogenic compositions. The polypeptides can be used in the manufacture
 CC of medicaments for treating or preventing infection due to *Neisseria*
 CC bacteria (e.g. meningitis and septicemia), to detect the presence of
 CC *Neisseria* bacteria, or to raise antibodies. They may also be used to
 CC screen for agonists or antagonists, which may themselves have use as
 CC antibacterial agents. The polynucleotides of the invention may also be
 CC used in gene therapy protocols. (Updated on 12-SEP-2003 to standardise OS
 CC field)
 XX Sequence 217 AA;
 SQ

Query Match 5.9%; Score 7; DB 3; Length 217;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLAALAA 26
 DB 14 VLAALAA 20
 |||||

RESULT 803
 AAU72926
 ID AAU72926 standard; protein; 217 AA.
 XX
 AC AAU72926;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE *Neisseria meningitidis* virulence protein #16.
 XX
 KW Meningitis; virulence; gene; antibacterial; vaccine; veterinary;
 KW infection; Gram-negative bacteria; antimicrobial.
 XX
 OS *Neisseria meningitidis*.
 XX
 FN WO200185772-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 08-MAY-2001; 2001WO-GB002003.
 XX
 PR 08-MAY-2000; 2000GB-00011108.

XX (MICR-) MICROSCIENCE LTD.
 XX Tang C;
 XX WPI; 2002-066593/09.
 DR N-PSDB; AA597211.
 XX New peptide encoded by operon including virulence genes of *Neisseria*
 PT meningitidis, useful as vaccine component for treating or preventing
 PT meningitis and for identifying antimicrobial drug.
 XX Claim 4; Page 75-76; 423pp; English.
 XX The invention relates to a peptide (I) encoded by an operon (II) of
 CC *Neisseria meningitidis* including virulence genes, or a related molecule
 CC having a 40% sequence similarity at the peptide or nucleotide level in a
 CC Gram-negative bacterium, or its functional fragment, for therapeutic or
 CC diagnostic use. (I) and (II) are useful in the manufacture of a
 CC medicament for treating or preventing a condition (e.g., meningitis)
 CC associated with infection by *Neisseria* or Gram-negative bacteria. The
 CC product is useful for veterinary treatment and in a screening assay for
 CC the identification of an antimicrobial drug. The vaccines have
 CC prophylactic applications. AAU72911-AAU73014 represent *N. meningitidis*
 CC virulence proteins of the invention
 XX Sequence 217 AA;
 SQ

Query Match 5.9%; Score 7; DB 5; Length 217;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLAALAA 26
 DB 14 VLAALAA 20
 |||||

RESULT 804
 ABP77518
 ID ABP77518 standard; protein; 217 AA.
 XX
 AC ABP77518;
 XX
 DT 07-MAR-2003 (first entry)
 XX
 DE *N. gonorrhoeae* amino acid sequence SEQ ID 1566.
 XX
 KW Antibacterial; infection; vaccine; gene therapy.
 XX
 OS *Neisseria gonorrhoeae*.
 XX
 FN WO200279243-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 12-FEB-2002; 2002WO-IB002069.
 XX
 PR 12-FEB-2001; 2001GB-00003424.
 XX
 XX (CHIR-) CHIRON SPA.
 XX Fontana MR, Pizza M, Masignani V, Monaci E;
 XX WPI; 2003-058415/05.
 DR N-PSDB; ABZ38488.
 XX
 XX New protein from *Neisseria gonorrhoeae*, useful for the manufacture of a
 PT medicament for treating or preventing *N. gonorrhoeae* infection.
 XX Disclosure; Page 300; 815pp; English.
 XX The present invention relates to proteins from *Neisseria gonorrhoeae*.
 CC Also disclosed are the nucleic acid molecules encoding the proteins and

CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
SQ Sequence 217 AA;

Query Match 5.9%; Score 7; DB 6; Length 217;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 14 VLAALAA 20
|||||
|

RESULT 805
ID AAR50078 standard; protein; 222 AA.
XX
AC AAR50078;
XX
DT 27-AUG-2003 (revised)
DT 24-NOV-1994 (first entry)
XX
DE NANBH virus antigenic fragment #10.
XX
KW Antigen; structural; non-structural; non A non B hepatitis virus; NANBHV;
KW NANBH; patient; plasma; diagnosis; detection; carrier; SB.
XX
OS Non-A.
OS non-B hepatitis virus.
XX
PN JP06070778-A.
XX
PD 15-MAR-1994.
XX
PF 01-JUN-1993; 93JP-00156087.
XX
PR 10-JUL-1992; 92JP-00207391.
XX
PA (TOKR-) ZH TOKYO RINSHO IGAKU SOGO KENKYUSHO.
PA (SANW) SANWA KAGAKU KENKYUSHO CO.
PA (TOFU) TONEN CORP.
PA (KOKU-) KOKUSAI SHIYAKU KK.
XX
DR WPI, 1994-128677/16.
DR N-PSDB; AAQ58823.
XX
PT Nucleic acid fragment coding non-A non-B hepatitis virus antigen - useful
PT in diagnosis of NANB patient and detection of virus carrier.
XX
PS Claim 20; Page 27-28; 37pp; Japanese.
XX
CC The sequences given in AAR50068 and AAR50070-82 represent antigens of
CC structural and non-structural regions of non A non B hepatitis virus
CC (NANBHV). The cDNA encoding these sequences were derived from the plasma
CC of a NANBH patient by recombinant DNA techniques. These fragments are
CC useful for the diagnosis of NANBH patients and the detection of NANBH
CC carriers. (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ Sequence 222 AA;

Query Match 5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 121 PDKEVLY 127
|||||
|

RESULT 807
ID AAB76730 standard; protein; 222 AA.
XX
AC AAB76730;
XX
DT 11-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MCT protein SEQ ID NO:442.
XX
KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;

RESULT 806
ID AAR49655 standard; protein; 222 AA.
XX
AC AAR49655;
XX
DT 25-MAR-2003 (revised)
DT 02-AUG-1994 (first entry)
XX
DE HCV peptide C14-8.
XX
KW HCV; hepatitis C virus group-I; hepatitis C virus group-II; grouping;
KW diagnosis; amplification; primer; RT-PCR;
KW reverse transcription polymerase chain reaction.
XX
OS Hepatitis C virus.
XX
PN EP586065-A2.
XX
PD 09-MAR-1994.
XX
PF 16-JUL-1993; 93EP-00305591.
XX
PR 16-JUL-1992; 92JP-00212061.
PR 30-OCT-1992; 92JP-00316834.
PR 30-OCT-1992; 92JP-00316835.
PR 30-APR-1993; 93JP-00104754.
XX
PA (TOFU) TONEN CORP.
XX
PI Hasegawa A, Maki N, Yagi S, Kashiwakuma T, Yamaguchi K;
PI Ikeguchi N, Kobayashi T, Senoo C;
XX
DR WPI, 1994-076364/10.
DR N-PSDB; AAQ58486.
XX
PT New antigenic peptide(s) specific for hepatitis C antibodies - allowing
PT differentiation between viral gps. I and II, useful for diagnosis and
PT grouping.
XX
PS Disclosure; Page 25-26; 35pp; English.
XX
CC DNA of sequence AAQ58486 encodes peptide C14-8 (AAR49655). This peptide
CC is a precursor of HCV C14-2-2 (given in AAR49654) and is obtained by PCR
CC cloning of HCV genes using the primers given in AAQ58480-83. (Updated on
CC 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PI
CC field.)
XX
SQ Sequence 222 AA;

Query Match 5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 121 PDKEVLY 127
|||||
|

RESULT 807
ID AAB76730 standard; protein; 222 AA.
XX
AC AAB76730;
XX
DT 11-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MCT protein SEQ ID NO:442.
XX
KW Corynebacterium glutamicum; brevibacterium lactofermentum; MCT;
KW membrane construction and membrane transport protein; petroleum spill;
KW hydrocarbon degradation; gram positive aerobic bacterium; marker;

KW identification; microorganism; fine chemical production; transformation;
XX genome mapping; genetic engineering.

OS Corynebacterium glutamicum.

XX WO200100805-A2.

XX 04-JAN-2001.

XX 23-JUN-2000; 2000WO-IB000926.

XX 25-JUN-1999; 99US-0141031P.

XX 08-JUL-1999; 99DE-01031454.

XX 08-JUL-1999; 99DE-01031478.

XX 08-JUL-1999; 99DE-01031563.

XX 09-JUL-1999; 99DE-01032122.

XX 09-JUL-1999; 99DE-01032124.

XX 09-JUL-1999; 99DE-01032125.

XX 09-JUL-1999; 99DE-01032128.

XX 09-JUL-1999; 99DE-01032180.

XX 09-JUL-1999; 99DE-01032182.

XX 09-JUL-1999; 99DE-01032190.

XX 09-JUL-1999; 99DE-01032191.

XX 09-JUL-1999; 99DE-01032209.

XX 09-JUL-1999; 99DE-01032212.

XX 09-JUL-1999; 99DE-01032227.

XX 09-JUL-1999; 99DE-01032228.

XX 09-JUL-1999; 99DE-01032229.

XX 09-JUL-1999; 99DE-01032230.

XX 14-JUL-1999; 99DE-01032927.

XX 14-JUL-1999; 99DE-01033005.

XX 14-JUL-1999; 99DE-01033006.

XX 27-AUG-1999; 99DE-01040832.

XX 27-AUG-1999; 99DE-01040833.

XX 31-AUG-1999; 99DE-01041378.

XX 31-AUG-1999; 99DE-01041379.

XX 31-AUG-1999; 99DE-01041395.

XX 03-SEP-1999; 99DE-01042077.

XX 03-SEP-1999; 99DE-01042078.

XX 03-SEP-1999; 99DE-01042079.

XX 03-SEP-1999; 99DE-01042088.

XX (BADI) BASF AG.

XX Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;

XX WPI; 2001-071486/08.

XX N-PSDB; RAB67963.

XX Corynebacterium glutamicum nucleic acids encoding membrane construction
XX and membrane transport proteins or their portions, useful for typing or
XX identifying C. glutamicum or related bacteria, and as markers for
XX transformation.

XX Claim 20; Page 789; 1119pp; English.

XX AAF67743 to AAF68080 encode the Corynebacterium glutamicum membrane
XX construction and membrane transport (MCT) proteins given in AAF676510 to
XX AAF676847. The MCT nucleic acids and proteins are useful in the
XX identification of microorganisms which can be used to produce fine
XX chemicals, for modulating fine chemical production in C. glutamicum or
XX related bacteria (e.g. Brevibacterium lactofermentum), the typing or
XX identification of C. glutamicum or related bacteria, as reference points
XX for mapping C. glutamicum genome, and as markers for transformation.
XX AAF68082 and AAF68082 represent sequencing primers which are used in an
XX example from the present invention

XX Sequence 222 AA;

Query Match

Best Local Similarity 5.9%; Score 7; DB 4; Length 222;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRAT 94

DB 58 GLLQRAT 64

RESULT 808

AAG90702

ID AAG90702 standard; protein; 222 AA.

XX AAG90702;

XX 26-SEP-2001 (first entry)

XX C glutamicum protein fragment SEQ ID NO: 4456.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;

XX organic acid synthesis.

XX Corynebacterium glutamicum.

XX EP1108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; RAB65921.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
XX mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 4456; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein
XX sequences from the Corynebacterium glutamicum glutamicum. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of corynebacterium, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX Corynebacterium, and identifying a homologue of a gene derived from
XX corynebacterium. Corynebacterium bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a protein described in the
XX exemplification of the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from the European Patent Office

XX Sequence 222 AA;

Query Match

Best Local Similarity 5.9%; Score 7; DB 4; Length 222;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRAT 94

DB 58 GLLQRAT 64

CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 222 AA;

Query Match 5.9%; Score 7; DB 6; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
DB 206 GGVLAAL 212
|||||

RESULT 811
ADJ81683
ID ADJ81683 standard; protein; 222 AA.
XX
XX AC ADJ81683;

DT 06-MAY-2004 (first entry)

DE Non-A-non-B hepatitis antigen amino acid sequence 10.

KW non-A-non-B type hepatitis virus antigen; recombinant technique;
KW hepatitis C virus infection.

XX Hepatitis C virus.

PN JP2004000151-A.

XX 08-JAN-2004.

PF 24-FEB-2003; 2003JP-00046384.

PR 10-JUL-1992; 92JP-00207391.

PR 01-JUN-1993; 93JP-00156087.

XX (KOKU-) KOKUSAI SHIYAKU KK.

XX WPI; 2004-085214/09.

XX N-PSDB; ADJ81669.

PT Novel nucleic acid fragment which codes for non-A-non-B type hepatitis
PT virus antigen, useful for diagnosing hepatitis C virus infection in a
PT patient.

XX Disclosure; Page 45-46; 59pp; Japanese.

CC This invention relates to a novel nucleic acid fragment containing a
CC nucleotide sequence which codes for non-A-non-B type hepatitis virus
CC antigen which has a fully defined sequence of 273 or 330 amino acids as
CC given in the specification. The invention is useful for producing non-A-
CC non-B type hepatitis virus antigen by recombinant techniques. The
CC invention may therefore be useful for diagnosing hepatitis C virus
CC infection in a patient and thus helping in prevention of the disease. The
CC invention allows effective detection of non-A-non-B hepatitis patients.
CC The present sequence is that of a non-A-non-B type hepatitis virus
CC antigen encoded by a DNA sequence of the invention. Note: This sequence
CC appears embedded within a cDNA sequence in the sequence listing of the
CC specification.

XX Sequence 222 AA;

Query Match 5.9%; Score 7; DB 8; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 121 PDKEVLY 127
|||||

RESULT 812

ADJ95020
ID ADJ95020 standard; protein; 227 AA.
XX
XX ADR95020;

DT 16-DEC-2004 (first entry)

DE Novel *S. pneumoniae* protein sequence, SEQ ID 3655.

KW Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
KW bacterial infection.

XX Streptococcus pneumoniae.

XX US6800744-B1.

XX 05-OCT-2004.

PF 30-JUN-1998; 98US-00107433.

PR 02-JUL-1997; 97US-0051553P.

PR 12-MAY-1998; 98US-0085131P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Doucette-Stamm LA, Bush D;

XX WPI; 2004-697205/68.

XX N-PSDB; ADR92417.

PT New isolated nucleic acid encoding a *Streptococcus pneumoniae*
PT polypeptide, useful for diagnosing, preventing and/or treating
PT pathological conditions resulting from the bacterial infection.

XX Disclosure; SEQ ID NO 3655; 151pp; English.

CC The invention relates to an isolated nucleic acid comprising a sequence
CC encoding a *Streptococcus pneumoniae* ADR91366polypeptide, or its
CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR95682,
CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
CC hybridisable under high stringency conditions to the nucleotide sequences.
CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
CC Also included are a recombinant expression vector comprising the isolated
CC nucleic acid cited above operably linked to a transcription regulatory
CC element, a cell comprising the recombinant expression vector and a probe
CC comprising at least 20 consecutive nucleotides of the nucleotide
CC sequences as cited above. The methods and compositions of the present
CC invention are useful for the diagnosis, prevention and/or treatment of
CC pathological conditions resulting from bacterial infection by
CC *Streptococcus pneumoniae* e.g. pneumonia, bacteraemia, meningitis and
CC otitis media. The present sequence is one of the 2603 disclosed *S.*
CC *pneumoniae* protein sequences. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=6800744B1.

XX Sequence 227 AA;

Query Match 5.9%; Score 7; DB 8; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24

Db 211 GGVLAAL 217
|||||||
Query Match 5.9%; Score 7; DB 9; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAL 24
Db 211 GGVLAAL 217
|||||||
RESULT 813
AEA58890
ID AEA58890 standard; protein; 227 AA.
XX AEA58890;
XX 25-AUG-2005 (first entry)
XX Streptococcus pneumoniae ORF amino acid sequence SEQ ID NO:3655.
DE bacterial infection; Streptococcus pneumoniae infection; antibacterial;
XX vaccine.
KW Streptococcus pneumoniae.
XX US2005136404-A1.
XX 23-JUN-2005.
XX 10-JUL-2003; 2003US-00617320.
XX 02-JUL-1997; 97US-0051553P.
XX 12-MAY-1998; 98US-0085131P.
XX 30-JUN-1998; 98US-00107433.
XX (DOUC/) DOUCETTE-STAMM L A.
XX (BUSH/) BUSH D.
XX Doucette-Stamm LA, Bush D;
PI WPI; 2005-477576/48.
XX N-PSDB; AEA56287.
XX New isolated nucleic acid molecules and encoded polypeptides useful for
diagnosing, preventing or treating bacterial infections, particularly
Streptococcus pneumoniae infection.
XX Claim 5; SEQ ID NO 3655; 144pp; English.
XX The invention relates to an isolated nucleic acid molecule for detecting,
preventing or treating pathological conditions resulting from bacterial
infection. The isolated nucleic acid comprises: (a) any of the 2603
nucleotide sequences of AEA55236 to AEA57838; (b) a nucleotide sequence
encoding a Streptococcus pneumoniae polypeptide comprising any of the
2603 amino acid sequences of AEA57839 to AEA60441; or (c) a nucleotide
sequence of at least 8 nucleotides in length, where the sequence is
hybridizable to a nucleic acid having any of the nucleotide sequences in
(a). Also described: (1) a recombinant expression vector comprising the
above nucleic acid operably linked to a transcription regulatory element;
(2) a cell comprising the recombinant expression vector; (3) producing an
S. pneumoniae polypeptide; (4) a probe comprising a nucleotide sequence
consisting of at least 8 nucleotides of any of AEA55236 to AEA57838; (5)
treating a subject for S. pneumoniae infection; (6) a recombinant or
substantially pure preparation of an S. pneumoniae polypeptide or its
fragment, where the polypeptide is selected from AEA57839 to AEA60441;
(7) a vaccine composition for preventing or treating an S. pneumoniae
infection, comprising an amount of the above nucleic acid or polypeptide;
(8) detecting the presence of a Streptococcus nucleic acid in a sample;
(9) a computer readable medium having recorded the nucleotide sequences
of AEA55236 to AEA57838; (10) a computer based system for identifying
fragments of the Streptococcus genome of commercial importance. The
composition and methods are useful for diagnosing, preventing or treating
bacterial infections, particularly S. pneumoniae infection. The present
sequence represents a S. pneumoniae ORF amino acid sequence from the
present invention. Note - The sequence data for this patent did not form
part of the printed specification, but was obtained in electronic format
directly from the USPTO web site.
XX Sequence 227 AA;
XX

Query Match 5.9%; Score 7; DB 9; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAL 24
Db 211 GGVLAAL 217
|||||||
RESULT 814
ADX77181
ID ADX77181 standard; protein; 230 AA.
XX ADX77181;
XX 21-APR-2005 (first entry)
XX Plant full length insert polypeptide seqid 46547.
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content.
XX Unidentified.
XX OS
XX US2004034888-A1.
XX 19-FEB-2004.
XX 28-APR-2003; 2003US-00425114.
XX 06-MAY-1999; 99US-00304517.
XX 05-NOV-2001; 2001US-00985678.
XX (LIU/) LIU J.
XX (ZHOU/) ZHOU Y.
XX (KOVA/) KOVALIC D K.
XX (SCRE/) SCREEN S E.
XX (TABA/) TABASKA J E.
XX (CAOY/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
pests, for conferring increased resistance to plant disease, or for
improving yield.
XX Claim 1; SEQ ID NO 46547; 15pp; English.
XX The invention describes a recombinant DNA construct comprising a
polynucleotide consisting of a sequence encoding an amino acid sequence
available in electronic form from the US patent office at
ftp.segdata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
of the invention are also useful in physical arrays of molecules and as
plant breeding markers. The recombinant DNA construct is useful for
improving plant tolerance to cold, heat, drought, herbicides, extreme
osmotic conditions, pathogens or pests, for manipulating growth rate in
plant cells by modification of the cell cycle pathway, for conferring
increased resistance to plant disease, for producing galactomannan,
lignin or plant growth regulators, for increasing the rate of homologous
recombination in plants, for improving yield by modification of
photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
or by providing improved plant growth and development under at least one
stress condition or for modifying seed oil or protein yield and/or

CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.

XX Sequence 230 AA;

Query Match 5.9%; Score 7; DB 8; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 VLLGGVL 21
DB 110 VLLGGVL 116
|||||

RESULT 815
ADJ79593
ID ADJ79593 standard; protein; 245 AA.
XX
XX
AC ADJ79593;
XX
XX 06-MAY-2004 (first entry)
XX
XX Hepatitis C virus polypeptide protein HCVb SeqID 102.
XX
XX bacterial host system; polypeptide; vaccine; cytotoxic T-cell lymphocyte;
KW HCV; infectious mononucleosis; nasopharyngeal carcinoma; virucidal;
KW cytostatic; hepatotropic; antiinflammatory; anti-HIV; hepatitis C virus.
XX
XX Human immunodeficiency virus.
XX
XX WO2004007556-A1.
XX
XX 22-JAN-2004.
XX
XX 14-JUL-2003; 2003WO-AU000910.
XX
XX 12-JUL-2002; 2002AU-00950193.
XX
XX (CSL-) CSL LTD.
XX
XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.
XX
XX Webb EA, Schoofs P;
XX
XX WPI; 2004-122896/12.
XX
XX Designing a candidate polypeptide for expression in a host, useful for
PT preventing or treating e.g. HIV, comprises identifying hydrophobic
PT peptide sequences in the polypeptide and arranging or re-locating the
PT peptide sequences.
XX
XX Disclosure; SEQ ID NO 102; 98pp; English.
XX
XX This invention relates to a novel method for designing heterologous
CC polypeptides comprising a proportion of hydrophobic amino acids that
CC increases the probability of the protein being efficiently expressed in a
CC bacterial host system. Specifically, it refers to arranging or re-
CC locating a hydrophobic peptide sequence within the protein of interest in
CC order to generate a candidate protein with reduced amplitude in
CC hydrophobicity and/ or length of any hydrophobic region. The present
CC invention describes using these polypeptides for use in a polypeptide
CC vaccine that is capable of eliciting a cytotoxic T-cell lymphocyte immune
CC response. As such, compositions can be used for the prevention or
CC treatment of diseases associated with the Epstein-Barr virus (EBV) such
CC as infectious mononucleosis or nasopharyngeal carcinoma, as well as
CC hepatitis C virus, cytomegalovirus or HIV. Accordingly, they exhibit
CC virucidal, cytostatic, hepatotropic, antiinflammatory and anti-HIV
CC activities. This polypeptide sequence is an HCV polypeptide protein
CC (HCVb) of the invention.
XX
XX Sequence 245 AA;

Query Match 5.9%; Score 7; DB 8; Length 245;

Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 GGVLAAL 24
DB 131 GGVLAAL 137
|||||

RESULT 816
ADC38707
ID ADC38707 standard; protein; 250 AA.
XX
XX AC ADC38707;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human secreted protein #32.
XX
XX immune disorder; severe combined immunodeficiency; SCID;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW rheumatoid arthritis; allergic reaction; asthma; myeloid cell deficiency;
KW lymphoid cell deficiency; osteoporosis; osteoarthritis;
KW peripheral nervous system disease; peripheral neuropathy;
KW Alzheimer's disease; Parkinson's disease; coagulation disorder;
KW inflammatory disease; systemic inflammatory response syndrome; SIRS;
KW ischaemia-reperfusion injury; Crohn's disease; anaphylaxis;
KW hypersensitivity; regeneration; neural cell proliferation; fertility;
KW tumour; chemokine; human; secreted protein.
XX
XX Homo sapiens.
OS
XX US2002193567-A1.
XX
XX 19-DEC-2002.
XX
XX 02-APR-2002; 2002US-00114893.
XX
XX 11-AUG-1995; 95US-00514014.
XX 05-APR-1996; 96US-00628364.
XX 19-APR-1996; 96US-00635311.
XX 07-JUN-1996; 96US-00659224.
XX 17-JUN-1996; 96US-00664596.
XX 09-JUL-1996; 96US-00677231.
XX 26-JUL-1996; 96US-00686878.
XX 23-AUG-1996; 96US-00701819.
XX 27-SEP-1996; 96US-00721488.
XX 27-SEP-1996; 96US-00721798.
XX 27-SEP-1996; 96US-00721923.
XX 27-SEP-1996; 96US-00721926.
XX 25-OCT-1996; 96US-00738367.
XX 30-OCT-1996; 96US-00739775.
XX 13-JAN-1997; 97US-00783395.
XX 10-APR-1997; 97US-00833823.
XX 02-JUN-1997; 97US-00867677.
XX 05-SEP-1997; 97US-00924838.
XX 06-OCT-1999; 99US-00413232.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Bowman MR, Spaulding V, Carlin-Duckett M;
PI Kelleher K;
XX
XX WPI; 2003-657236/62.
XX N-PSDB; ADC38706.
XX
XX Proteins AZ3021 encoded by clone AZ3021 from human adult colon, and
BD12716 encoded by clone BD12716 from human fetal kidney cDNA library,
PT useful for treating e.g. multiple sclerosis and rheumatoid arthritis.
XX
XX Disclosure; SEQ ID NO 65; 412pp; English.
XX
XX The invention relates to a protein comprising fully defined AZ302 1
CC

CC protein or BD127 1 6 protein. The polynucleotides are useful for
 CC expressing recombinant proteins for analysis and are also useful as
 CC chromosome markers or tags to identify chromosomes or to map related gene
 CC positions. The proteins are useful as amino acid supplement, carbon
 CC source, nitrogen source and carbohydrate source. The proteins are useful
 CC for treating various immune deficiencies and disorders (e.g. severe
 CC combined immunodeficiency (SCID)), autoimmune disorders (e.g. multiple
 CC sclerosis, systemic lupus erythematosus, rheumatoid arthritis), allergic
 CC reactions (e.g. asthma), myeloid or lymphoid cell deficiencies,
 CC osteoporosis or osteoarthritis, peripheral nervous system diseases (e.g.
 CC peripheral neuropathy, Alzheimer's disease, Parkinson's disease),
 CC coagulation disorders, inflammatory diseases (e.g. systemic inflammatory
 CC response syndrome (SIRS), ischaemia-reperfusion injury, Crohn's disease),
 CC anaphylaxis and hypersensitivity. Proteins are also useful for inducing
 CC tumour immunity, for inducing bone, cartilage, tendon, ligament and/or
 CC nerve growth or regeneration, for proliferating neural cells and for
 CC regenerating nerve and brain tissue, for inducing fertility and for
 CC inhibiting tumour growth. Proteins are also useful as chemokine for
 CC mammalian cells (e.g., monocytes, fibroblasts, neutrophils), and also
 CC useful as inhibitors of receptor/ligand interactions. The present
 CC sequence represents the amino acid sequence of a human secreted protein.
 CC
 SQ Sequence 250 AA;

Query Match 5.9%; Score 7; DB 7; Length 250;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

QY 18 GGVLAAL 24
 DB 203 GGVLAAL 209
 |||||

RESULT 817
 ABU22837
 ID ABU22837 standard; protein; 255 AA.
 AC ABU22837;
 XX
 XX 19-JUN-2003 (first entry)
 XX
 XX Protein encoded by Prokaryotic essential gene #8364.
 XX
 XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 XX Burkholderia mallei.
 XX
 XX WO200277193-A2.
 XX
 XX 03-OCT-2002.
 XX
 XX 21-MAR-2002; 2002WO-US009107.
 XX
 XX 21-MAR-2001; 2001US-00815242.
 XX
 XX 06-SEP-2001; 2001US-00948993.
 XX
 XX 25-OCT-2001; 2001US-0342923P.
 XX
 XX 08-FEB-2002; 2002US-00072851.
 XX
 XX 06-MAR-2002; 2002US-0362699P.
 XX
 XX (ELIT-) ELITRA PHARM INC.
 XX
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 XX WPI; 2003-029926/02.
 XX
 XX N-P8DB; ACA26707.
 XX
 XX New antisense nucleic acids, useful for identifying proteins or screening
 XX for homologous nucleic acids required for cellular proliferation to
 XX isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 50761; 1766pp; English.

CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 255 AA;

Query Match 5.9%; Score 7; DB 6; Length 255;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0;

QY 20 VLAALAA 26
 DB 244 VLAALAA 250
 |||||

RESULT 818
 ADA49405
 ID ADA49405 standard; protein; 255 AA.
 XX
 XX ADA49405;
 XX
 XX 20-NOV-2003 (first entry)
 XX
 XX Multi-epitope construct #5.
 XX
 XX multi-epitope; immunogenic; epitope; major histocompatibility complex;
 XX MHC class I; MHC class II; junctional epitope.
 XX
 XX Synthetic.
 XX
 XX Hepatitis C virus.
 XX
 XX US2002119127-A1.
 XX
 XX 29-AUG-2002.
 XX
 XX 27-JUN-2001; 2001US-00894018.
 XX
 XX 28-DEC-1999; 99US-0173390P.
 XX
 XX 28-DEC-2000; 2000WO-US035568.
 XX
 XX 16-APR-2001; 2001US-0284221P.
 XX
 XX (SETT)/ SETTE A.
 XX (CHES)/ CHESNUT R.
 XX (LIVI)/ LIVINGSTON B D.
 XX (BAKE)/ BAKER D M.


```
XX SQ Sequence 255 AA;
Query Match 5.9%; Score 7; DB 9; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 101 GGVLAAL 107

RESULT 821
AAR20611
ID AAR20611 standard; protein; 256 AA.
XX AC AAR20611;
XX DT 27-AUG-2003 (revised)
DT 06-MAY-1992 (first entry)
XX XX
XX C10-13 NANBH-specific antigen polypeptide.
XX Non-A non-B hepatitis virus; recombinant; detection.
XX OS Non-A.
XX OS non-B hepatitis virus.
XX PN EP468657-A.
XX XX
XX PD 29-JAN-1992.
XX PF 09-JUL-1990; 90JP-00180889.
XX PR 09-JUL-1990; 90JP-00180889.
XX PR 30-NOV-1990; 90JP-00339589.
XX PR 20-DEC-1990; 90JP-00413844.
XX XX
XX PA (TOFU) TONEN CORP.
XX PI Maki N, Yamaguchi K, Toyoshima A, Kohara M;
XX WPI; 1992-034390/05.
XX DR N-PSDB; AAR20619.
XX PT Non-A, non-B hepatitis-specific antigen polypeptide - for detection of
PT hepatitis virus gene or antibody directed against the virus.
XX PS Disclosure; Fig 3; 78pp; English.
XX CC The amino acid sequence is that of a non-A non-B (NANB) hepatitis
CC specific antigen polypeptide, which may be recombinantly produced. It is
CC encoded by a fragment of DNA clone C10-13. It can be used to prepare
CC antibodies which can be used to detect NANB hepatitis with extremely high
CC accuracy. It can also be used to detect anti-NANB hepatitis antibodies.
CC See also AAR20609-R20615 and AAR20713-R20723. (Updated on 27-AUG-2003 to
CC correct OS field.)
XX SQ Sequence 256 AA;
Query Match 5.9%; Score 7; DB 2; Length 256;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 191 PDKEVLY 197

RESULT 822
AAB43436
ID AAB43436 standard; protein; 260 AA.
XX
```

```
AC AAB43436;
XX 08-FEB-2001 (first entry)
DT Human cancer associated protein sequence SEQ ID NO:881.
DE Human; cancer associated gene; cancer antigen; detection; cancer;
DE diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
DE antidiabetic; antiaesthetic; antirheumatic; antithyroid; antiviral;
DE dermatological; antithyroid; antiallergic; antibacterial; cardiant;
DE vasotropic; antipneumatic; antidiabetic; thrombolytic; coagulant; nontropic;
DE immune disorder; haematopoietic cell disorder; gene therapy; inflammation;
DE allergic reaction; graft versus host disease; organ rejection;
DE haemostatic; thrombolytic; cardiovascular disorder; infection;
DE neurological disease; drug screening.
XX OS Homo sapiens.
XX PN WC20005350-A1.
XX PD 21-SEP-2000.
XX PF 08-MAR-2000; 2000WC-US005882.
XX PR 12-MAR-1999; 99US-0124270P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX DR WPI; 2000-587533/55.
XX DR N-PSDB; AAC77645.
XX PT Novel isolated nucleic acids comprising sequences encoding peptides
XX useful for treating or diagnosing e.g. cancer.
XX PS Claim 11; Page 1436-1437; 2352pp; English.
XX CC AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; vulnery; immunomodulator;
CC antidiabetic; antiaesthetic; antirheumatic; antithyroid; antiviral;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; coagulant;
CC dermatological; vasotropic; antipneumatic; antidiabetic; thrombolytic;
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC disorders, allergic reactions, graft versus host disease and organ
CC rejection, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX SQ Sequence 260 AA;
Query Match 5.9%; Score 7; DB 3; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 QQQAVIE 101
Db 51 QQQAVIE 57

RESULT 823
```

AAU64862
ID AAU64862 standard; protein; 260 AA.
XX
AC AAU64862;
XX
DT 27-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #25758.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
FN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PP 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
XX
PR 02-JUN-2000; 2000US-0208841P.
XX
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
XX
DR N-PSDB; AAS59654.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Example 1; SEQ ID NO 26057; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 260 AA;

Query Match 5.9%; Score 7; DB 4; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 173 VLAALAA 179
|||||||
|||||||

RESULT 824

AAE19895
ID AAE19895 standard; protein; 260 AA.
XX
AC AAE19895;
XX
DT 18-JUN-2002 (first entry)
XX
DE Hepatitis C virus (HCV) NS4B protein.
XX
KW Hepatitis C virus; HCV infection; virucide; fungicide; antibacterial;
XX cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
OS Hepatitis C virus.
XX
FN WO200213855-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001808.
XX
PR 17-AUG-2000; 2000US-0225767P.
XX
PR 29-AUG-2000; 2000US-0229175P.
XX
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX
PI Sallberg M, Hultgren C;
XX
DR WPI; 2002-241837/29.
XX
PT Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
PS Claim 11; Page 77-78; 120pp; English.
XX
CC The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is HCV NS4B protein
XX
SQ Sequence 260 AA;

Query Match 5.9%; Score 7; DB 5; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNWQKLE 112
DB 41 TNWQKLE 47
|||||||
|||||||

RESULT 825
ABM61381
ID ABM61381 standard; protein; 260 AA.
XX
AC ABM61381;
XX
DT 20-OCT-2003 (first entry)
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #26057.
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX
OS Propionibacterium acnes.
XX
FN WO2003033515-A1.
XX

```
PD 24-APR-2003.
XX
XX PF 11-OCT-2002; 2002WO-US032727.
XX
XX PR 15-OCT-2001; 2001US-00978825.
XX
XX PA (CORI-) CORIYA CORP.
XX
XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX PI Barth B, Vallieue-Douglas J;
XX
XX XX WPI; 2003-381789/36.
XX
XX DR N-PSDB; ACF64583.
XX
XX
XX PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX PS Example 1; SEQ ID NO 26057; 1481pp; English.
XX
XX CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX CC encoding a Propionibacterium acnes protein. The invention also relates to
XX CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX CC immunogenic fragments of P. acnes polypeptides. The invention
XX CC additionally encompasses expression vectors and host cells comprising a
XX CC polynucleotide of the invention; antibodies against polypeptides of the
XX CC invention; fusion proteins comprising a polypeptide of the invention; a
XX CC method for stimulating an immune response specific for a P. acnes
XX CC polypeptide and an isolated T cell population comprising T cells prepared
XX CC via this method; a vaccine composition (comprising P. acnes polypeptides,
XX CC polynucleotides, antibodies, fusion proteins, T cell populations, or
XX CC antigen-presenting cells that express the polypeptide); a method and kit
XX CC for detecting or determining the presence or absence of P. acnes in a
XX CC patient; and a method for inhibiting the development of P. acnes in a
XX CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations or antigen-presenting cells that express the
XX CC polypeptides are useful for diagnosing, preventing or treating acne
XX CC vulgaris, or for stimulating an immune response specific for a P. acnes
XX CC protein. The polynucleotides can also be used as probes or primers for
XX CC nucleic acid hybridisation. The vaccine composition is useful for the
XX CC stimulation of an immune response against P. acnes, or for treating acne,
XX CC and the kit is useful for performing a diagnostic assay. The present
XX CC sequence represents a polypeptide predicted to be encoded by an ORF (open
XX CC reading frame) contained within the P. acnes polynucleotides of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 260 AA;

Query Match 5.9%; Score 7; DB 6; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAALAA 26
Db 173 VLAAALAA 179
|||||

RESULT 826
ABW00346
ID ABW00346 standard; protein; 260 AA.
XX
XX AC ABW00346;
XX
XX DT 15-JAN-2004 (first entry)
XX
XX DE Hepatitis C virus NS4B protein.
XX
XX KW Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide.

Hepatitis C virus.
US2002136740-A1.
26-SEP-2002.
15-AUG-2001; 2001US-00929955.
17-AUG-2000; 2000US-0225767P.
29-AUG-2000; 2000US-0229175P.
(SALL/) SALLBERG M.
(HULT/) HULTGREN C.
Sallberg M, Hultgren C;
WPI; 2003-764978/72.
Vaccine compositions for treating and preventing disease, preferably
hepatitis C virus infection, comprises ribavirin and antigen that has
epitope present in hepatitis C virus.
Claim 11; Page 40-41; Opp; English.
The invention relates to a composition comprising ribavirin and an
antigen, where the antigen is derived from a hepatitis virus. The vaccine
is useful in enhancing the immune response to a hepatitis C antigen where
the composition is delivered to an animal identified as requiring an
enhanced immune response. The vaccine is useful in the treatment and
prevention of hepatitis C infection. The present sequence is Hepatitis C
virus NS4B protein
Sequence 260 AA;

Query Match 5.9%; Score 7; DB 7; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
Db 41 TNWQKLE 47
|||||

RESULT 827
ABU43166
ID ABU43166 standard; protein; 262 AA.
XX
XX AC ABU43166;
XX
XX DT 19-JUN-2003 (first entry)
XX
XX DE Protein encoded by Prokaryotic essential gene #28693.
XX
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX OS Staphylococcus epidermidis.
XX
XX PN WO200277183-A2.
XX
XX PD 03-OCT-2002.
XX
XX PF 21-MAR-2002; 2002WO-US009107.
XX
XX PR 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
```

XX WPI; 2003-029926/02.
DR N-PSDB; ACA47036.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 71090; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 262 AA;
Query Match 5.9%; Score 7; DB 6; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 LGGVLA 22
|||||
Db 191 LGGVLA 197
RESULT 828
ABU19953
ID ABU19953 standard; protein; 262 AA.
XX
AC ABU19953;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #5480.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
OS *Borrelia cepacia*.
XX
FN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX

PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA23823.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 47877; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 262 AA;
Query Match 5.9%; Score 7; DB 6; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLALAA 26
|||||
Db 251 VLALAA 257
RESULT 829
ABP39361
ID ABP39361 standard; protein; 263 AA.
XX
AC ABP39361;
XX
DT 24-JUL-2002 (first entry)
XX
DE *Staphylococcus epidermidis* ORF amino acid sequence SEQ ID NO:4206.
XX

19-APR-2004; 2004WO-US012047.
18-APR-2003; 2003US-0463708P.
18-APR-2003; 2003US-0463732P.
02-MAY-2003; 2003US-0467199P.
02-MAY-2003; 2003US-0467230P.
19-MAY-2003; 2003US-0471306P.
19-MAY-2003; 2003US-0471336P.
08-JUL-2003; 2003US-0485223P.
08-JUL-2003; 2003US-0485224P.
14-JUL-2003; 2003US-0486446P.
14-JUL-2003; 2003US-0486480P.
08-AUG-2003; 2003US-0493573P.
08-AUG-2003; 2003US-0493577P.
08-SEP-2003; 2003US-0505059P.
(FIVE-) FIVE PRIME THERAPEUTICS INC.
PA
PI Lee E, Hestir K, Chu K, Masuoka L, Williams LT;
XX
XX WPI; 2004-775861/76.
DR N-PSDB; ADU01949.
XX
XX New first nucleic acid molecule comprising a polynucleotide sequence,
PT given in the specification, useful in preparing a composition for
PT diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
XX
XX Claim 14; SEQ ID NO 1148; 291pp; English.
XX
XX The invention describes a new first nucleic acid molecule comprising a
CC polynucleotide sequence given in the specification. Also described are:
CC an animal injected with the nucleic acid molecule; a second nucleic acid
CC molecule comprising a second polynucleotide sequence that is at least
CC about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
CC that hybridises to the first polynucleotide sequence under high
CC stringency conditions; a vector comprising the nucleic acid molecule and
CC a promoter that drives the expression of the nucleic acid molecule; a
CC host cell transformed, transfected, transduced or infected with the
CC nucleic acid molecule; a nucleic acid composition comprising a carrier or
CC a buffer and one or more compositions comprising the nucleic acid
CC molecule, vector or host cell; a substantially purified polypeptide; an
CC animal injected with the polypeptide; a polypeptide composition
CC comprising the polypeptide molecule and a carrier or buffer; a cell
CC culture medium comprising the polypeptide or transfected cells
CC transfected with the polynucleotide; making a transformed, transfected,
CC transduced, or infected host cell; synthesising Nanodiscs simultaneously
CC and for synthesising a series of simultaneously-synthesised Nanodiscs
CC sequentially utilising a dynamic system; preparing a hydrophobic protein
CC for determination of crystal structure; immunising a non-human animal;
CC screening for modulators of hydrophobic protein activity; a diagnostic
CC kit; determining the presence of the nucleic acid molecule or its
CC complement; determining the presence of an antibody to the polypeptide in
CC a sample; an antibody specifically recognising, binding to or modulating
CC the biological activity of at least one polypeptide encoded by a nucleic
CC acid molecule or its biologically active fragment; an antibody
CC composition comprising the antibody and a carrier; a bacteriophage, where
CC the antibody is displayed on the bacteriophage; a bacterial cell
CC comprising the bacteriophage; a non-human animal injected with the
CC antibody composition; a host cell that secretes the antibody; making an
CC antibody; diagnosing a disease, disorder, syndrome, or condition
CC comprising cancer, or proliferative, inflammatory, immune, metabolic,
CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
CC conditions in a patient; a modulator composition comprising a modulator
CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a
CC subject; an isolated modified cell comprising at least one first
CC heterologous nucleic acid molecule, where the first heterologous nucleic
CC acid molecule comprises a first polynucleotide sequence that encodes a
CC first polypeptide; a non-human animal deficient in the polypeptide or
CC that over-expresses the polypeptide; isolated tissues derived from the
CC non-human animal; and one or more cells derived from the non-human
CC animal. The nucleic acid is useful in preparing a composition for
CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.

CC This is the amino acid sequence of a novel human polypeptide of the
CC invention.
XX

SQ Sequence 266 AA;

Query Match 5.9%; Score 7; DB 8; Length 266;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCVWIVG 38
|||||
Db 142 GCVWIVG 148

RESULT 832

AAW92815

ID AAW92815 standard; protein; 269 AA.

XX

AC AAW92815;

XX

DT 10-MAY-1999 (first entry)

XX

DE HCV NS4B protein.

XX

KW NS4B protein; ATPase; Flaviviridae; antiviral compound; Dengue virus;
KW therapeutic agent; yellow fever virus; pestiviruses; swine fever;
KW bovine viral diarrhoea virus.

XX

OS Hepatitis C virus.

XX

PN WO9901582-A1.

XX

PD 14-JAN-1999.

XX

PF 01-JUL-1998; 98WO-US013790.

XX

PR 02-JUL-1997; 97US-0051582P.

XX

PA (SMIK) SMITHKLINE BEECHAM CORP.

XX

PI Delvecchio A, Zhong W;

XX

WPI; 1999-106080/09.

XX

PT Detecting Hepatitis C Virus NS4B protein modulators - useful to treat

XX infection with viruses of the Flaviviridae family.

XX

PS Claim 6; Page 21-22; 27pp; English.

XX

This sequence is used in a novel method for identifying a compound that
alters activity of Hepatitis C Virus (HCV) NS4B protein. HCV NS4B
modulators can be used as antiviral compounds and as therapeutic agents
to treat viruses of the Flaviviridae family, including HCV, yellow fever
virus, Dengue viruses types 1-4, and pestiviruses such as bovine viral
diarrhoea virus and classic swine fever. Treatment of an HCV-infected
human with a HCV1 NS4B antagonist, and treatment of an HCV-infected
mammal with a HCV NS4B agonist is claimed

SQ Sequence 269 AA;

Query Match 5.9%; Score 7; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNWQKLE 112
|||||
Db 41 TNWQKLE 47

RESULT 833

AAG27984

ID AAG27984 standard; protein; 273 AA.

XX

AC AAG27984;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 33032.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; Genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137588P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 26-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149910P.
PR 25-AUG-1999; 99US-0150566P.

PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 30-AUG-1999; 99US-0151080P.
 PR 31-AUG-1999; 99US-0151303P.
 PR 01-SEP-1999; 99US-0151438P.
 PR 07-SEP-1999; 99US-0151930P.
 PR 10-SEP-1999; 99US-0152363P.
 PR 13-SEP-1999; 99US-0153070P.
 PR 15-SEP-1999; 99US-0153758P.
 PR 16-SEP-1999; 99US-0154018P.
 PR 20-SEP-1999; 99US-0154039P.
 PR 22-SEP-1999; 99US-0154779P.
 PR 23-SEP-1999; 99US-0155139P.
 PR 24-SEP-1999; 99US-0155486P.
 PR 28-SEP-1999; 99US-0155659P.
 PR 29-SEP-1999; 99US-0156458P.
 PR 04-OCT-1999; 99US-0156596P.
 PR 05-OCT-1999; 99US-0157117P.
 PR 06-OCT-1999; 99US-0157531P.
 PR 07-OCT-1999; 99US-0157865P.
 PR 08-OCT-1999; 99US-0158029P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 14-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 21-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161992P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 5.9%; Score 7; DB 3; Length 273;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
 |||||
 Db 216 VLLGGVL 222

RESULT 834
 ADR04184
 ID ADR04184 standard; protein; 273 AA.

XX AC ADR04184;
 XX DT 07-OCT-2004 (first entry)
 XX DE E faecium Muri protein SEQ ID NO: 48.

XX Muri; antibacterial; vulnery; auditory; antiinflammatory; vaccine;
 KW

KW glutamate racemase; enzyme; protein co-ordinate data.
 XX Enterococcus faecium.
 OS WO2004061097-A2.
 PN 22-JUL-2004.
 XX 08-DEC-2003; 2003WO-US038977.
 XX 20-DEC-2002; 2002US-0435087P.
 PR 20-DEC-2002; 2002US-0435167P.
 PR 20-DEC-2002; 2002US-0435272P.
 PR 20-DEC-2002; 2002US-0435527P.
 XX (ASTR) ASTRAZENECA AB.
 PA (ASTR) ASTRAZENECA PHARM LP.
 XX Anderson M, Fisher SL, Folmer RHA, Kern G, Lundqvist RT;
 PI Newton DT, Xue Y;
 XX WPI; 2004-553367/53.
 DR N-PSDB; ADR04183.
 XX New crystal of glutamate racemase, useful for treating bacterial
 PT infection, e.g. endocarditis, wound infection, respiratory tract
 PT infection, including sinusitis, otitis media, bacterial meningitis,
 PT peptic ulcer, or gastric MALT.
 XX Claim 21; SEQ ID NO 48; 1125pp; English.
 XX The present invention provides the crystal structure of glutamate
 CC racemase (Muri) complexed with an inhibitor and a substrate. The Muri
 CC inhibitor is useful for treating bacterial infection. The crystal,
 CC methods of identifying its inhibitors, compositions and vaccines
 CC containing the inhibitors are useful for treating bacterial infection,
 CC e.g. endocarditis, wound infection, respiratory tract infection,
 CC including sinusitis, otitis media, bacterial meningitis, peptic ulcer, or
 CC gastric MALT. The present sequence is a Muri protein.
 XX Sequence 273 AA;
 SQ

Query Match 5.9%; Score 7; DB 8; Length 273;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIELGGK 45
 |||||
 Db 267 HIELGGK 273

RESULT 835
 AARS4203
 ID AARS4203 standard; protein; 277 AA.

XX AC AARS4203;
 XX 25-MAR-2003 (revised)

DT 18-NOV-1994 (first entry)
 XX snab gene product involved in streptogramin biosynthetic pathway.

XX Antibiotic; streptogramin; snaA; snaB; snaC; biosynthesis; enzyme;
 KW biosynthetic pathway; Streptomyces pristinaespiralis.
 XX Streptomyces pristinaespiralis.

XX FR2696189-A1.
 XX 01-APR-1994.

XX 25-SEP-1992; 92FR-00011441.
 XX

PR 25-SEP-1992; 92FR-00011441.
XX (RHON) RHONE POULENC RORER SA.
PA
XX Blanc V, Blanche F, Crouzet J, Jacques N, Lacroix P, Thibaut D;
XX Zagorec M;
PI
XX WPI; 1994-128286/16.
XX N-PSDB; AAQ64203.
DR
XX
XX DNA involved in streptogramin antibiotic biosynthesis - for prodn. or bio
PT -conversion of streptogramin(s) or prodn. of streptogramin intermediates,
PT derivs. or hybrid antibiotics.
XX
XX Claim 21; Page 52-53; 83pp; French.
XX
XX The snabB gene product is involved in the biosynthesis of streptogramins,
CC antibiotics active against Gram-positive bacteria. The identification of
CC the sequences encoding the enzymes involved in the biosynthetic pathway
CC means that they can be isolated and manipulated. Mutant microorganisms in
CC which a step in the streptogramin biosynthetic pathway is blocked can be
CC cultured to produce streptogramin intermediates, which may later be
CC converted to streptogramin derivatives. Recombinant cells may also be
CC used for the bioconversion of streptogramins from one form to another or
CC for the production of hybrid antibiotics. (Updated on 25-MAR-2003 to
CC correct FN field.)
XX
XX SQ Sequence 277 AA;

Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
Db 240 VLALAA 246
|||||

RESULT 836
AAG27983
ID AAG27983 standard; protein; 277 AA.
AC
AC AAG27983;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 33031.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-012180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.

23-APR-1999; 99US-0130891P.
28-APR-1999; 99US-0131449P.
30-APR-1999; 99US-0132048P.
04-MAY-1999; 99US-0132407P.
05-MAY-1999; 99US-0132484P.
06-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
07-MAY-1999; 99US-0132487P.
11-MAY-1999; 99US-0132863P.
14-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
18-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
16-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
29-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.
19-JUL-1999; 99US-0144332P.
19-JUL-1999; 99US-0144333P.
19-JUL-1999; 99US-0144334P.
19-JUL-1999; 99US-0144335P.
20-JUL-1999; 99US-0144352P.
20-JUL-1999; 99US-0144632P.

```
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156598P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157533P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.

PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 26-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 5.9%; Score 7; DB 3; Length 277;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
Db 220 VLLGGVL 226
|||||

RESULT 837
ABU29809
ID ABU29809 standard; protein; 277 AA.
XX
AC ABU29809;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #15336.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Enterococcus faecium.
XX
PN WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
WPI; 2003-029926/02.
DR N-PSDB; ACA33679.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 57733; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
the 6213 antisense sequences given in the specification where expression
```

CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: the sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 277 AA;
 SQ
 Query Match 5.9%; Score 7; DB 6; Length 277;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELGGK 45
 Db 271 HIELGGK 277
 |||||

RESULT 838
 ADX96062
 ID ADX96062 standard; protein; 288 AA.
 XX AC ADX96062;
 XX AC

DT 21-APR-2005 (first entry)

DE Plant full length insert polypeptide seqid 58726.

XX plant protectant; plant growth regulant; gene therapy; plant;
 KW recombinant DNA construct; physical array; plant breeding marker;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
 KW extreme osmotic condition; pathogen tolerance; pest tolerance;
 KW growth rate; cell cycle pathway; disease resistance;
 KW galactomannan production; lignin production; plant growth regulator;
 KW yield; plant growth; plant development; seed oil; protein yield;
 KW protein content.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIUJ/) LIU J.

PA (ZHOU/) ZHOU Y.

PA (KOVA/) KOVALIC D K.
 PA (SCRE/) SCREEN S E.
 PA (TABA/) TABASKA J E.
 PA (CAOY/) CAO Y.

PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
 PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
 PT pests, for conferring increased resistance to plant disease, or for
 PT improving yield.

XX Claim 1; SEQ ID NO 58726; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
 CC polynucleotide consisting of a sequence encoding an amino acid sequence
 CC available in electronic form from the US patent office at
 CC ftp.segdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
 CC of the invention are also useful in physical arrays of molecules and as
 CC plant breeding markers. The recombinant DNA construct is useful for
 CC improving plant tolerance to cold, heat, drought, herbicides, extreme
 CC osmotic conditions, pathogens or pests, for manipulating growth rate in
 CC plant cells by modification of the cell cycle pathway, for conferring
 CC increased resistance to plant disease, for producing galactomannan,
 CC lignin or plant growth regulators, for increasing the rate of homologous
 CC recombination in plants, for improving yield by modification of
 CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
 CC or by providing improved plant growth and development under at least one
 CC stress condition or for modifying seed oil or protein yield and/or
 CC content. This is the amino acid sequence of a plant full length insert
 CC polypeptide that can be used in the recombinant DNA construct of the
 CC invention.

XX Sequence 288 AA;

Query Match 5.9%; Score 7; DB 8; Length 288;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQ 91
 |||||

Db 47 KVLGLLQ 53

RESULT 839

ADX08670

ID ADE08670 standard; protein; 294 AA.

XX AC ADE08670;

DT 29-JAN-2004 (first entry)

DE Novel protein (useful for identifying genetic disorders) #825.

XX novel gene; novel protein; tissue marker; molecular weight marker;
 KW chromosome marker; genetic disorder.

XX Unidentified.

XX WO2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002WO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

XX 11-DEC-2001; 2001US-0339453P.

XX 14-MAR-2002; 2002US-0365091P.

XX 14-MAR-2002; 2002US-0365384P.

XX 12-APR-2002; 2002US-0372381P.

XX 12-APR-2002; 2002US-0372615P.

PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI; 2003-569235/53.
DR N-PSDB; ADE07759.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
PS Claim 20; SEQ ID NO 1736; 1177pp; English.
XX
CC The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence represents a protein
CC of the invention.
XX
XX Sequence 294 AA;
SQ
Query Match 5.9%; Score 7; DB 7; Length 294;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GVLAAAL 25
Db 9 GVLAAAL 15
|||||
RESULT 840
ID AAR90027 standard; protein; 295 AA.
AC AAR90027;
XX
XX 22-JUL-1996 (first entry)
XX
XX Methionine aminopeptidase sequence.
XX
XX Methionine aminopeptidase; heat resistant.
XX
XX Pyrococcus furiosus.
XX
XX JP08009979-A.
XX
XX 16-JAN-1996.
XX
XX 05-JUL-1994; 94JP-00174930.
XX
XX 05-JUL-1994; 94JP-00174930.
XX
XX (TAKI) TAKARA SHUZO CO LTD.
XX
XX WPI; 1996-110277/12.
DR N-PSDB; AAT12495.
XX
XX Heat-resistant methionine aminopeptidase - useful for selective removal
PT of initiator methionine residue from N-terminus of recombinant proteins;
XX
XX Claim 3; Page 10; 13pp; Japanese.
PS
XX This sequence represents a heat resistant methionine aminopeptidase. This
CC protein keeps 95% or more of its activity upon treatment at 75 degrees C
CC for 60 minutes. The methionine aminopeptidase allows selective deletion
CC

CC of initiator methionine from the N-terminal of a precursor protein
CC produced by recombinant DNA technology. A large amount of high purity
CC heat resistant methionine aminopeptidase can be produced
XX
SQ Sequence 295 AA;
Query Match 5.9%; Score 7; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 ELGGKPA 47
Db 43 ELGGKPA 49
|||||
RESULT 841
ID ADN46674 standard; protein; 295 AA.
XX
XX ADN46674;
AC
XX
XX 01-JUL-2004 (first entry)
XX
XX Thermococcus kodakaraensis KOD1 protein sequence SeqID552.
XX
XX gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;
KW gene structure; gene function; enzyme activity; medicine;
KW forensic science; food; drug inspection; molecular biology; immunology.
XX
XX Thermococcus kodakaraensis.
OS
XX WO2004022736-A1.
PN
XX
XX 18-MAR-2004.
PD
XX
XX 29-AUG-2003; 2003WO-IB003597.
PF
XX
XX 30-AUG-2002; 2002JP-00319011.
PR
XX
XX (NTSC-) JAPAN SCI & TECHNOLOGY CORP.
PA
XX
XX Imanaka T, Atomi H;
PI
XX
XX WPI; 2004-257583/24.
DR
XX
XX Method for disrupting targeted gene in genome of organism particularly
PT thermotable bacterium and with genome chips for analysis, applicable in
PT studying gene structure and functions.
XX
XX Claim 9; SEQ ID NO 552; 598pp; Japanese.
PS
XX
XX This invention relates to a novel method for targeting disruption of an
CC arbitrary gene in a genome of an organism which comprises providing the
CC whole sequential data of the genome of such organism, selecting at least
CC 1 arbitrary region in the sequence, providing a vector that contains a
CC sequence homologous with the selected region and a marker gene.
CC transformation, and homologous recombination. The genome is preferably
CC the genome of a hyperthermostable archaeobacterium, particularly
CC Thermococcus kodakaraensis KOD1. The method is for targeting the
CC disruption of a gene in the genome of an organism, which is applicable in
CC studying gene structure and functions as well as enzyme activities of
CC encoded proteins and useful in medicine, forensic science, food or drug
CC inspection, molecular biology and immunology. With this method, the
CC disruption of a gene at an arbitrary position in a genome can be achieved
CC efficiently and reliably. The present sequence is that of a protein
CC encoded by the genome of Thermococcus kodakaraensis which was derived
CC using the method of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 295 AA;

Query Match 5.9%; Score 7; DB 8; Length 295;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47
Db 44 ELGGKPA 50
|||||

RESULT 842
AAG81135
ID AAG81135 standard; protein; 306 AA.
XX
AC AAG81135;
XX
DT 04-SEP-2001 (first entry)
XX
DE Mycobacterium tuberculosis potential drug target protein SEQ ID 186.
XX
KW Drug target; growth; organism viability; characterisation.
XX
OS Mycobacterium tuberculosis.
XX
PN WO200135317-A1.
XX
PD 17-MAY-2001.
XX
PF 13-NOV-2000; 2000WO-US031152.
XX
PR 12-NOV-1999; 99US-0165086P.
PR 12-NOV-1999; 99US-0165124P.
PR 01-FEB-2000; 2000US-0179531P.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Eisenberg D, Rotstein SH, Marcotte EM;
XX
DR WPI; 2001-329193/34.
DR N-PSDB; AAH51986.
XX
PT Identifying nucleotide or polypeptide sequence for use as drug target,
PT involves providing algorithm that analyzes a functional relationship
PT between nucleotide or polypeptide sequences, and comparing the sequences.
XX
PS Disclosure; Page 165; 207pp; English.
XX
SQ This invention relates to a method for identifying a nucleotide or
CC polypeptide sequence that may be a drug target, or essential for growth
CC or viability of an organism. Polynucleotide sequences AAH51947 - AAH52092
CC represent DNA encoding proteins AAG81096 - AAG81241, Mycobacterium
CC tuberculosis proteins which are potential drug targets. The DNA and
CC protein sequences are used to illustrate the method of the invention. The
CC method involves providing an unknown nucleotide or polypeptide sequences,
CC and comparing it to a number of sequences along with at least one
CC algorithm capable of analysing a functional relationship between
CC nucleotide and polypeptide sequences. The method is useful for
CC characterising the function of nucleic acids and polypeptides that may be
CC useful as a target for a drug or essential for the growth or viability of
CC an organism
XX

Query Match 5.9%; Score 7; DB 4; Length 306;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
Db 206 VLALAA 212
|||||

RESULT 843
AAB68498
ID AAB68498 standard; protein; 316 AA.
XX
AC AAB68498;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32286.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL12601.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 32286; 21pp + Sequence Listing; English.
XX
SQ The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AAB57737-
CC AAB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

Query Match 5.9%; Score 7; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 127 LLGGVLA 133
|||||

RESULT 844
AAY92506
ID AAY92506 standard; protein; 319 AA.
XX
AC AAY92506;
XX
DT 10-AUG-2000 (first entry)
XX
DE Human OXRE-3 with identity to lambda crystallin.
XX
KW OXRE-3; oxidoreductase; lambda crystallin; antiproliferative; anticancer;
KW hepatotropic; antiviral; antitasthmatic; nootropic; neuroprotectant;
KW antiparkinsonian's; antisclerotic; anxiolytic; antischizophrenic;
KW anti-inflammatory; antiarthritic.
XX

```
OS Homo sapiens.
XX Key Location/Qualifiers
XX Domain 9..283
XX FT /label= signature_sequence
XX FT 61
XX FT /note= "potential phosphorylation site"
XX FT 67
XX FT /note= "potential phosphorylation site"
XX FT 111
XX FT /note= "potential phosphorylation site"
XX FT 167
XX FT /note= "potential phosphorylation site"
XX FT 216
XX FT /note= "potential phosphorylation site"
XX FT 218
XX FT /note= "potential phosphorylation site"
XX FT 250
XX FT /note= "potential phosphorylation site"
XX WO200020604-A2.
XX 13-APR-2000.
XX 06-OCT-1999; 99WO-US023434.
XX 06-OCT-1999; 98US-0172227P.
XX 02-DEC-1998; 98US-0155202P.
XX 10-MAR-1999; 99US-0123911P.
XX (INCY-) INCYTE PHARM INC.
XX Lal P, Guegler KJ, Gorgone GH, Corley NC, Baughn MR, Tang YT;
XX Hillman JL, Bandman O, Azimzai Y, Au-Young J, Yue H, Lu DAM;
XX Yang J;
XX WPI; 2000-303785/26.
XX N-PSDB; AAA09377.
XX Purified polypeptide for treating or preventing disorders associated with
XX decreased expression or activity of oxidoreductase molecules.
XX Claim 1; Page 71-72; 97pp; English.
XX AAY92504-18 show oxidoreductases, designated OXRE-1 to -15. The
XX polypeptides are useful for treating or preventing a disorder associated
XX with decreased expression or activity of OXRE. Antagonists of OXRE are
XX useful for treating or preventing a disorder associated with increased
XX expression or activity of OXRE. The disorders include cell proliferative
XX disorders (cirrhosis, hepatitis), cancer (leukemia, melanoma),
XX hypopituitarism and hyperpituitarism, hypothyroidism and hyperthyroidism,
XX metabolic disorders (Addison's disease, cystic fibrosis), reproductive
XX disorders (infertility, ovulatory defects), neurological disorders
XX (Alzheimer's disease, Parkinson's disease, multiple sclerosis), mental
XX disorders (anxiety, schizophrenia), autoimmune/inflammatory disorders
XX (acquired immunodeficiency syndrome (AIDS), asthma, osteoarthritis), and
XX viral infections. The polynucleotides may be used in Southern or Northern
XX analysis, polymerase chain reaction (PCR), or in enzyme-linked
XX immunosorbent assays (ELISA)
XX Sequence 319 AA;
XX Query Match 5.9%; Score 7; DB 3; Length 319;
XX Best Local Similarity 100.0%; Pred. No. 4.3e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 32 GCWIVG 38
XX Db 7 GCWIVG 13
XX RESULT 845
XX AAB94861
```

```
ID AAB94861 standard; protein; 319 AA.
XX AC AAB94861;
XX DT 26-JUN-2001 (first entry)
XX DE Human protein sequence SEQ ID NO:16058.
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX Claim 8; SEQ ID NO 16058; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAHQ3166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX Sequence 319 AA;
XX Query Match 5.9%; Score 7; DB 4; Length 319;
XX Best Local Similarity 100.0%; Pred. No. 4.3e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 32 GCWIVG 38
XX Db 7 GCWIVG 13
XX RESULT 846
```

[illegible]

PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

 Query Match 5.9%; Score 7; DB 3; Length 321;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 15 VLLGGVL 21
 DB 264 VLLGGVL 270

 RESULT 847
 ADE08132
 ID ADE08132 standard; protein; 324 AA.
 XX AC
 AC ADE08132;
 XX
 DT 29-JAN-2004 (first entry)
 XX

DE Novel protein (useful for identifying genetic disorders) #287.
 XX novel gene; novel protein; tissue marker; molecular weight marker;
 KW chromosome marker; genetic disorder.
 XX Unidentified.
 XX WO2003054152-A2.
 XX 03-JUL-2003.
 XX 10-DEC-2002; 2002WO-US039555.
 XX 10-DEC-2001; 2001US-0339739P.
 PR 11-DEC-2001; 2001US-0339453P.
 PR 14-MAR-2002; 2002US-0365091P.
 PR 14-MAR-2002; 2002US-0365384P.
 PR 12-APR-2002; 2002US-0372381P.
 PR 12-APR-2002; 2002US-0372615P.
 PR 22-APR-2002; 2002US-00128558.
 PR 24-APR-2002; 2002US-0376045P.
 XX (HYSE-) HYSEQ INC.
 PA Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
 PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
 XX WPI; 2003-569235/53.
 DR N-PSDB; ADE07221.
 XX
 PT New polynucleotides, useful for expressing recombinant proteins for
 analysis, characterization or therapeutic use, or as markers for tissues
 in which the corresponding protein is preferentially expressed.
 XX Claim 20; SEQ ID NO 1198; 1177pp; English.
 PS The invention comprises the amino acid and coding sequences of novel
 proteins. The DNA and protein sequences of the invention are useful as:
 CC markers for tissues in which the corresponding protein is preferentially
 expressed; as molecular weight markers on gels; as chromosome markers or
 CC tags; to identify chromosomes or to map related gene positions; and to
 CC compare with endogenous DNA sequences in patients to identify potential
 genetic disorders. The present amino acid sequence represents a protein
 CC of the invention.
 XX
 SQ Sequence 324 AA;

 Query Match 5.9%; Score 7; DB 7; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 18 GGVLAAL 24
 DB 202 GGVLAAL 208

 RESULT 848
 ADN26699
 ID ADN26699 standard; protein; 335 AA.
 XX AC
 AC ADN26699;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #9352.
 XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polypeptide.
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 9352; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant
XX comprising the recombinant DNA construct and a method of producing a
XX transformed plant having an improved property. The plant is a crop plant
XX such as maize or soybean. The method of producing a transformed plant
XX having an improved property comprises transforming a plant with the
XX recombinant DNA construct and growing the transformed plant, where the
XX polynucleotide or polypeptide is useful for improving plant properties.
XX The recombinant DNA construct is useful for producing plants with
XX improved plant properties, e.g. improved cold, heat or drought tolerance,
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX increased resistance to plant disease, better growth rate by modification
XX of the cell cycle pathway with plant growth regulators, increased rate of
XX homologous recombination, modified seed oil or protein yield and/or
XX content, improved yield by modification of carbohydrate, nitrogen or
XX phosphorus use and/or uptake, by modification of photosynthesis or by
XX providing improved plant growth and development under at least one stress
XX condition, improved lignin production or improved galactomannan
XX production. This sequence represents a bacterial polypeptide used in the
XX scope of the invention. Note: The sequence data for this patent did not
XX form part of the printed specification but was obtained in electronic
XX format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 335 AA;
XX
XX Query Match 5.9%; Score 7; DB 8; Length 335;
XX Best Local Similarity 100.0%; Pred. No. 4.5e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 94 TQQQAVI 100
XX |||||
XX Db 102 TQQQAVI 108
XX
XX RESULT 849
XX ABM65633
XX ID ABM65633 standard; protein; 338 AA.
XX AC ABM65633;
XX DT 20-OCT-2003 (first entry)
XX XX
XX DE Propionibacterium acnes immunogenic polypeptide #30309.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine; immunogenic.
XX
XX Propionibacterium acnes.
XX WO2003033515-A1.
XX 24-APR-2003.
XX 11-OCT-2002; 2002WO-US032727.
XX 15-OCT-2001; 2001US-00978825.
XX (CORI-) CORIXA CORP.
XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
XX Barth B, Vallie-Douglas J;
XX WPI; 2003-381789/36.
XX
XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
XX or for stimulating an immune response specific for a P. acnes protein.
XX
XX Claim 7; SEQ ID NO 30309; 1481pp; English.
XX
XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
XX encoding a Propionibacterium acnes protein. The invention also relates to
XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
XX immunogenic fragments of P. acnes polypeptides. The invention
XX additionally encompasses expression vectors and host cells comprising a
XX polynucleotide of the invention; antibodies against polypeptides of the
XX invention; fusion proteins comprising a polypeptide of the invention; a
XX method for stimulating an immune response specific for a P. acnes
XX polypeptide and an isolated T cell population comprising T cells prepared
XX via this method; a vaccine composition (comprising P. acnes polypeptides,
XX polynucleotides, antibodies, fusion proteins, T cell populations, or
XX antigen-presenting cells that express the polypeptide); a method and kit
XX for detecting or determining the presence or absence of P. acnes in a
XX patient; and a method for inhibiting the development of P. acnes in a
XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
XX proteins, T cell populations or antigen-presenting cells that express the
XX polypeptides are useful for diagnosing, preventing or treating acne
XX vulgaris, or for stimulating an immune response specific for a P. acnes
XX protein. The polynucleotides can also be used as probes or primers for
XX nucleic acid hybridisation. The vaccine composition is useful for the
XX stimulation of an immune response against P. acnes, or for treating acne,
XX and the kit is useful for performing a diagnostic assay. The present
XX sequence represents a specifically claimed P. acnes polypeptide which is
XX thought to contain an immunogenic region. Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 338 AA;
XX
XX Query Match 5.9%; Score 7; DB 6; Length 338;
XX Best Local Similarity 100.0%; Pred. No. 4.6e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 20 VLAAALAA 26
XX |||||
XX Db 173 VLAAALAA 179
XX
XX RESULT 850
XX ADN26623
XX ID ADN26623 standard; protein; 341 AA.
XX XX
XX AC ADN26623;
XX XX

DT 02-DEC-2004 (first entry)
DE Bacterial polypeptide #9276.
XX
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
OS
XX US2003233675-A1.
XX
XX 18-DEC-2003.
XX
XX 20-FEB-2003; 2003US-00369493.
XX
XX 21-FEB-2002; 2002US-0360039P.
XX
XX (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 9276; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 341 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 341;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 94 TQOQAVI 100
DB 103 TQOQAVI 109

RESULT 851
ABB89956
ID ABB89956 standard; protein; 344 AA.
XX
AC ABB89956;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2332.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
XX WO200190304-A2.
XX
XX 29-NOV-2001.
XX
XX 18-MAY-2001; 2001WO-US016450.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
PI WPI; 2002-122018/16.
XX
XX N-PSDB; ABL90365.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX
XX Claim 11; SEQ ID NO 2332; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 344 AA;
SQ
Query Match 5.9%; Score 7; DB 5; Length 344;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 32 GCWVIVG 38
DB 7 GCWVIVG 13

RESULT 852
AAG64355
ID AAG64355 standard; protein; 351 AA.

```
XX AC AAG64355;
XX DT 01-OCT-2001 (first entry)
XX DE Human lambda crystallin.
XX KW Human; lambda crystallin.
XX OS Homo sapiens.
XX PN CN1275621-A.
XX PD 06-DEC-2000.
XX PF 26-MAY-1999; 99CN-00107102.
XX PR 26-MAY-1999; 99CN-00107102.
XX PA (UYFU-) UNIV FUDAN.
XX PI Yu L, Zhao Y, Zhang H;
XX DR WPI; 2001-211850/22.
XX DR N-PSDB; AAH73866.
XX PT Human lambda crystallin and its coding sequence, preparation process and
XX PR use.
XX PS Claim 2; Page 15 (Disclosure); 19pp; Chinese.
XX CC The present sequence is human lambda crystallin. This sequence is a
XX CC homologue of rabbit lambda crystallin
XX SQ Sequence 351 AA;

Query Match 5.9%; Score 7; DB 4; Length 351;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
Db 7 GCVVIVG 13

RESULT 853
ADU02842
ID ADU02842 standard; protein; 357 AA.
AC ADU02842;
XX DT 27-JAN-2005 (first entry)
XX DE Novel human polypeptide seqid 1309.
XX cytostatic; antipsoriatic; antiinflammatory; gene therapy; Nanodisc;
KW proliferative disorder; inflammatory disorder; immune disorder;
KW metabolic disorder; bone disorder; CNS disorder; cancer; psoriasis;
KW ulcerative colitis; human.
XX OS Homo sapiens.
XX PN WO2004039804-A2.
XX PD 04-NOV-2004.
XX PF 19-APR-2004; 2004WO-US012047.
XX PR 18-APR-2003; 2003US-0463708P.
XX PR 18-APR-2003; 2003US-0463732P.
XX PR 02-MAY-2003; 2003US-0467199P.
XX PR 02-MAY-2003; 2003US-0467230P.
XX PR 19-MAY-2003; 2003US-0471306P.

PR 19-MAY-2003; 2003US-0471336P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.
PR 08-SEP-2003; 2003US-0505059P.
XX (FIVE-) FIVE PRIME THERAPEUTICS INC.
XX PA
XX PI Lee E, Hestir K, Chu K, Masuoka L, Williams LT;
XX DR WPI; 2004-775861/76.
XX DR N-PSDB; ADU02110.
XX PT New first nucleic acid molecule comprising a polynucleotide sequence
XX PR given in the specification, useful in preparing a composition for
XX PT diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
XX PS Claim 14; SEQ ID NO 1309; 291pp; English.
XX CC The invention describes a new first nucleic acid molecule comprising a
XX CC polynucleotide sequence given in the specification. Also described are:
XX CC an animal injected with the nucleic acid molecule; a second nucleic acid
XX CC molecule comprising a second polynucleotide sequence that is at least
XX CC about 70, 80, 90 or 95% homologous to the first nucleic acid molecule or
XX CC that hybridises to the first polynucleotide sequence under high
XX CC stringency conditions; a vector comprising the nucleic acid molecule and
XX CC a promoter that drives the expression of the nucleic acid molecule; a
XX CC host cell transformed, transfected, transduced or infected with the
XX CC nucleic acid molecule; a nucleic acid composition comprising a carrier or
XX CC a buffer and one or more compositions comprising the nucleic acid
XX CC molecule, vector or host cell; a substantially purified polypeptide; an
XX CC animal injected with the polypeptide; a polypeptide composition
XX CC comprising the polypeptide molecule and a carrier or buffer; a cell
XX CC culture medium comprising the polypeptide or transfected cells
XX CC transfected with the polynucleotide; making a transformed, transfected,
XX CC transduced, or infected host cell; synthesising Nanodiscs simultaneously
XX CC and for synthesising a series of simultaneously-synthesised Nanodiscs
XX CC sequentially utilising a dynamic system; preparing a hydrophobic protein
XX CC for determination of crystal structure; immunising a non-human animal;
XX CC screening for modulators of hydrophobic protein activity; a diagnostic
XX CC kit; determining the presence of the nucleic acid molecule or its
XX CC complement; determining the presence of an antibody to the polypeptide in
XX CC a sample; an antibody specifically recognising, binding to or modulating
XX CC the biological activity of at least one polypeptide encoded by a nucleic
XX CC acid molecule or its biologically active fragment; an antibody
XX CC composition comprising the antibody and a carrier; a bacteriophage, where
XX CC the antibody is displayed on the bacteriophage; a bacterial cell
XX CC comprising the bacteriophage; a non-human animal injected with the
XX CC antibody composition; a host cell that secretes the antibody; making an
XX CC antibody; diagnosing a disease, disorder, syndrome, or condition
XX CC comprising cancer, or proliferative, inflammatory, immune, metabolic,
XX CC bone, CNS, genetic, bacterial and viral diseases, disorders, syndromes or
XX CC conditions in a patient; a modulator composition comprising a modulator
XX CC and a carrier; gene therapy; prophylactic or therapeutic treatment of a
XX CC subject; an isolated modified cell comprising at least one first
XX CC heterologous nucleic acid molecule, where the first heterologous nucleic
XX CC acid molecule comprises a first polynucleotide sequence that encodes a
XX CC first polypeptide; a non-human animal deficient in the polypeptide or
XX CC that over-expresses the polypeptide; isolated tissues derived from the
XX CC non-human animal; and one or more cells derived from the non-human
XX CC animal. The nucleic acid is useful in preparing a composition for
XX CC diagnosing or treating e.g., cancer, psoriasis or ulcerative colitis.
XX CC This is the amino acid sequence of a novel human polypeptide of the
XX CC invention.
XX SQ Sequence 357 AA;

Query Match 5.9%; Score 7; DB 8; Length 357;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 20 VLAALAA 26
Db 247 VLAALAA 253

RESULT 854
ADS28712
ID ADS28712 standard; protein; 359 AA.
XX AC
XX ADS28712;
DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #17745.
DE
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
OS
XX US2003233675-A1.
FN
XX 18-DEC-2003.
PD
XX 20-FEB-2003; 2003US-00369493.
PF
XX 21-FEB-2002; 2002US-0360039P.
PR
XX (CAO Y).
PA (HINKLE G J.
PA (SLATY) SLATER S C.
PA (CHEN X).
PA (GOLD) GOLDMAN B S.
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI WPI; 2004-061375/06.
DR
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
PS Claim 1; SEQ ID NO 17745; 122pp; English.

The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 359 AA;
Query Match 5.9%; Score 7; DB 8; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 94 TOQQA VI 100
Db 111 TOQQA VI 117

RESULT 855
ABU23531
ID ABU23531 standard; protein; 361 AA.
XX
AC ABU23531;
XX
DT 19-JUN-2003 (first entry)
XX Protein encoded by Prokaryotic essential gene #9058.
DE
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
OS Bordetella pertussis.
XX WO200277183-A2.
FN
XX 03-OCT-2002.
PD
XX 21-MAR-2002; 2002WO-US009107.
PF
XX 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
WPI; 2003-029926/02.
DR N-PSDB; ACA27401.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 51455; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of the 6213 antisense sequences given in the specification where expression of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 361 AA;

Query Match 5.9%; Score 7; DB 6; Length 361;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVIAA 23
Db 325 LGGVIAA 331
|||||

RESULT 856
AAE09651
ID AAE09651 standard; protein; 363 AA.
XX
AC AAE09651;
DT 19-NOV-2001 (first entry)
XX
DE Human.gene 13 encoded lipid metabolism protein HTUN173, SEQ ID NO:45.
XX
KW Human; lipid metabolism protein; nootropic; neuroprotective; cardiant;
KW cerebroprotective; hepatotropic; antidiabetic; ophthalmic; nephrotropic;
KW immune disorder; autoimmune disease; rheumatoid arthritis; glossitis;
KW systemic lupus erythematosus; conjunctivitis; inflammatory disorder;
KW respiratory disorder; asthma; allergy; CNS disorder; Alzheimer's disease;
KW Parkinson's disease; atherosclerosis; cardiovascular disorder; cancer;
KW coronary disease; familial hypercholesterolaemia; hyperlipidaemia;
KW haematopoietic disorder; hypolipidaemia; lipodosis; Gaucher's disease;
KW Tay-sach's disease; mental retardation; gene therapy; antisense therapy.
XX
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX
XX Misc-difference 337 /label= Unknown
XX /note= "Encoded by MYA"
XX
XX Misc-difference 338 /label= Unknown
XX /note= "Encoded by ASC"
XX
XX Misc-difference 340 /label= Unknown
XX /note= "Encoded by ARA"
XX
XX Misc-difference 360 /label= Unknown
XX /note= "Encoded by GNA"
XX
XX
XX WO200155203-A1.
XX
XX
XX 02-AUG-2001.
XX
XX
XX 17-JAN-2001; 2001WO-US001327.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216800P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233083P.
PR 14-SEP-2000; 2000US-0233084P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.

PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 14-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764866.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Ruben SM, Barash SC;
WPI; 2003-801167/75.
N-PSDB; AAD62062.
New nucleic acid molecules and polypeptides for diagnosing, preventing or treating disorders associated with aberrant expression of the polypeptide, e.g. neural or cardiovascular disorders, and in chromosome identification.
Claim 11; Page 202-203; 284pp; English.
AAD62050-AAD62071 represent cDNAs corresponding to 22 human secreted protein genes and ABW01075-ABW01096 represent the proteins they encode.
AAD62072-AAD62091 represent human genomic DNAs. The invention are useful in diagnosing, preventing, treating and ameliorating immune disorders

Query Match 5.9%; Score 7; DB 7; Length 363;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWIVG 38
Db 26 GCWIVG 32

RESULT 858
AAY40499
ID AAY40499 standard; protein; 366 AA.
XX
AC AAY40499;
XX
DT 03-DEC-1999 (first entry)
XX
DE Synechocystis yael polypeptide.
XX
KW Bacterial polypeptide; yael family; antibiotic; gram negative bacteria;
KW gram positive bacteria; therapeutic; medicament; antibiotic screening;
KW bacterial mediated disorder.
XX
OS Synechocystis sp.
XX
PW WO9947553-A2.
XX
PD 23-SEP-1999.
XX
PF 18-MAR-1999; 99WO-GB000850.
XX
PR 18-MAR-1998; 98GB-00005792.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Arigoni F, Edgerton MD, Lofrer H, Peitsch MC;
XX
DR WPI; 1999-562058/47.
XX
PT New bacterial polypeptides designated as yael family polypeptides, useful
PT for identification of novel broad-spectrum antibiotics.
XX
PS Claim 1; Fig 1; 96pp; English.
XX

The invention provides isolated bacterial polypeptides belonging to a family designated the yael family. The yael family polypeptides are defined by one of the following: (i) a high-scoring segment pair (HSP) score of greater than or equal to 250 when compared with one of the specific yael family member sequences given in the specification using the known BLAST search algorithm (Altschul, S.F. et al, J. Mol. Biol. (1990) 215:403-10) with a BLOSUM62 scoring matrix (Proteins (1993) 17:49-61); or (ii) identified by profile based searches using known MAST technique to yield a p-value of less than 1×10^{-30} and position-dependent scoring matrices given in the specification for yael family members; or (iii) comprising at least two amino acid sequences from (a) or sequence of (b): (a) h-e-(tlyf)-g-h-(lf)-x-x-a; n-l-x-p-x-l-d-g; and (pl)-(lv)-g-g-(yf)-(cv); and (b) h-e-(tlvmfyw)-g-h-(lf)-x(2)-a-x(30,40)-(pl)-(lv)-g-g-(yf)-(cv). Where: letters designate amino acid one letter codes; letters in square brackets denote alternative single amino acids; x is any amino acid; and numbers in brackets indicate the number of residues at that position. The polypeptides are useful to screen for agents with antibiotic activity. They are especially useful in the development of new broad-spectrum antibiotics, since they are conserved across, and believed to be essential for the viability of, taxonomically diverse bacteria, including both gram negative and gram positive bacteria. Antagonists identified by such screens may be administered therapeutically to treat bacterial infections and used to manufacture medicaments for such treatment. The polynucleotides encoding the polypeptides are useful in antibiotic screening as above, as probes for other members of the gene family or in antisense therapy to block or reduce polypeptide expression to treat bacterial mediated disorders. Sequences AAY40492-501 represent yael family polypeptides of different bacterial species

SQ Sequence 366 AA;
Query Match 5.9%; Score 7; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 3 VLAALAA 9

RESULT 859
AAB19718
ID AAB19718 standard; protein; 367 AA.
XX
AC AAB19718;
XX
DT 19-FEB-2001 (first entry)
XX
DE Arabidopsis thaliana SSE1 protein.
XX
KW SSE1; shrunken seed gene; storage reserve; storage protein; oil body;
KW transgenic plant.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FH Region 38..53
FT /note= "hydrophobic region"
FT Region 109..124
FT /note= "hydrophobic region"
FT Region 134..180
FT /note= "hydrophilic region"
FT Region 201..227
FT /note= "hydrophilic region"
FT Region 238..257
FT /note= "hydrophobic region"
FT Region 337..355
FT /note= "hydrophobic region"
XX
PW WO200061735-A1.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US009192.
XX
PR 08-APR-1999; 99US-0128651P.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Lin Y, Sun L, Nguyen LV, Goodman HM;
XX
DR WPI; 2000-679483/66.
DR N-PSDB; AAA88782.
XX
PT Novel shrunken seed gene useful for producing transgenic plants having
PT altered production of food storage reserve material, intracellular
PT transport of storage protein and formation of protein or oil bodies.
XX
PS Claim 2; Page 58; 64pp; English.
XX

The present sequence is that of Arabidopsis thaliana SSE1 (shrunken seed) protein, as deduced from isolated cDNA (see AAA88782). SSE1 protein, when expressed in a cell of a plant, modifies or alters the production of a food storage reserve material (e.g. protein, lipid or carbohydrate storage reserve), facilitates the intracellular transport of a storage protein, or facilitates the formation of protein or oil bodies. The invention provides a transgenic plant for plant cell, plant tissue, plant organ or plant component) which includes a recombinant SSE1 transgene that modifies the production of food storage reserves, thereby increasing nutritional value. An antisense construct is useful for modifying desiccation tolerance

SQ Sequence 367 AA;
Query Match 5.9%; Score 7; DB 3; Length 367;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 QQAVIEP 102
| | | | |
Db 209 QQAVIEP 215
RESULT 860
ADC35637
ID ADC35637 standard; protein; 367 AA.
XX AC ADC35637;
XX 18-DEC-2003 (first entry)
XX Arabidopsis thaliana SSE1 (shrunk seed 1) protein sequence.
XX SSE1; shrunk seed 1; flowering plant; seed; seedling growth; food;
KW feed; industrial purpose; SSE1; seed storage protein; lipid deposition;
KW starch deposition; endosperm; oil deposition; protein content;
KW oil content; storage organ formation; grain; corn.
XX Arabidopsis thaliana.
XX US2003084475-A1.
XX 01-MAY-2003.
XX 09-OCT-2002; 2002US-00268441.
XX 30-SEP-1999; 99US-0157209P.
XX 28-SEP-2000; 2000US-00672607.
XX (CAHO)/ CAHOON E B.
XX (COUG)/ COUGHLAN S J.
XX (HELE)/ HELENTJARIIS T G.
XX (JUNG)/ JUNG R.
XX (LICP)/ LI C P.
XX (NICH)/ NICHOLS S E.
XX (RIPP)/ RIPP K G.
XX (ZHEN)/ ZHENG P.
XX Cahoon EB, Coughlan SJ, Helentjaris TG, Jung R, Li CP;
PI Nichols SE, Ripp KG, Zheng P;
DR WPI; 2003-755127/71.
XX New isolated SSE1 polynucleotide for producing plants capable of
PT partitioning photosynthate to produce seed with improved functional
PT properties for use in specific food and non-food industrial applications.
XX Disclosure; SEQ ID NO 15; 38pp; English.
XX This invention relates to novel isolated nucleic acids which encode SSE1
CC (shrunk seed 1) homologues. Seeds of flowering plants contain proteins,
CC starches and oils in a balance suitable to support seedling growth of the
CC next generation. Users of seeds for food, feed or industrial purposes
CC often desire modifications in quality or quantity of these components.
CC SSE1 was cloned from Arabidopsis thaliana and is thought to be involved
CC in the pathway leading to seed storage protein and lipid deposition. If
CC this pathway is inactivated, seed starch deposition proceeds by default.
CC Altered expression of SSE1 can increase the starch deposition in the
CC endosperm, or increase the oil deposition in the embryo, increase the
CC protein content in the seed, or increase the oil and protein content in
CC the seed. The sequences of the invention may also be used to modulate
CC storage organ formation in the seed of a plant. Hence the invention may
CC be used to improve the food, feed and/or industrial value of grain. The
CC present sequence is the amino acid sequence of the Arabidopsis thaliana
CC SSE1 protein, which was used to identify the SSE1 homologues of the

CC invention.
XX SQ Sequence 367 AA;
Query Match 5.9%; Score 7; DB 7; Length 367;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 QQAVIEP 102
| | | | |
Db 209 QQAVIEP 215
RESULT 861
ABG29922
ID ABG29922 standard; protein; 379 AA.
XX AC ABG29922;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #29913.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX N-PSDB; AAS94109.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 60281; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG0377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 379 AA;

Query Match 5.9%; Score 7; DB 4; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 96 QOAVIEP 102
 |||||
 Db 191 QOAVIEP 197

RESULT 862
 AEB14376
 ID AEB14376 standard; protein; 389 AA.
 XX
 AC AEB14376;
 XX
 DT 22-SEP-2005 (first entry)
 XX
 DE Plant lipid metabolism protein SEQ ID NO 54.
 XX
 KW lipid metabolism protein; LMP; transgenic plant; plant; seed storage;
 KW crop improvement.
 XX
 OS core eudicotyledons.
 XX
 FN WO2005063995-A2.
 PD
 PD 14-JUL-2005.
 XX
 PP 22-DEC-2004; 2004WO-IB004251.
 XX
 PR 23-DEC-2003; 2003US-0532751P.
 XX
 PA (BADI) BASF PLANT SCI GMBH.
 XX
 PI Oswald O, Bauer J, Zank T;
 XX
 DR WPI; 2005-506658/51.
 DR N-PSDB; AEB14375.
 XX
 PT Novel isolated lipid metabolism protein (LMP) nucleic acid, useful for
 PT producing transgenic plant having modified level of seed storage compound
 PT e.g. lipids, fatty acids, starches and seed storage proteins.
 XX
 PS Claim 1; SEQ ID NO 54; 115pp; English.
 XX
 CC The invention relates to an isolated lipid metabolism protein (LMP)
 CC nucleic acid. The LMP or the nucleic acid encoding it is useful for
 CC producing a transgenic plant having a modified level of a seed storage
 CC compound. The LMP is useful for modulating the level of a seed storage
 CC compound e.g. lipids, fatty acids, starches and seed storage proteins, in
 CC a plant. The present sequence represents the amino acid sequence of a
 CC plant lipid metabolism protein.
 XX
 SQ Sequence 389 AA;

Query Match 5.9%; Score 7; DB 9; Length 389;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 VLLGGVL 21
 |||||
 Db 264 VLLGGVL 270

RESULT 863
 AA37377
 ID AA37377 standard; protein; 396 AA.
 XX
 AC AA37377;
 XX
 DT 27-AUG-2003 (revised)

DT 11-MAR-1998 (first entry)
 XX
 DE Hepatitis C virus chimeric peptide antigen.
 XX
 KW Hepatitis C virus; HCV; chimeric; antigen; detection; core region;
 KW epitope; NS3; NS4; infection.
 XX
 OS Synthetic.
 OS Hepatitis C virus.
 OS Chimeric.
 XX
 PN JP09278794-A.
 XX
 PD 28-OCT-1997.
 XX
 PF 10-FEB-1997; 97JP-00027015.
 XX
 PR 09-FEB-1996; 96JP-00024045.
 XX
 PA (TOFU) TONEN CORP.
 XX
 DR WPI; 1998-022248/03.
 DR N-PSDB; AAT97238.
 XX
 PT New chimaeric peptide antigen derived from hepatitis C virus protein -
 PT useful for detecting HCV infections.
 XX
 PS Claim 3; Page 26-28; 30pp; Japanese.
 XX
 CC The present sequence represents a Hepatitis C virus (HCV) chimeric
 CC peptide antigen which comprises at least 2 peptide epitope regions from
 CC the HCV polypeptide core region, 2 peptide epitope regions from the NS3
 CC region and at least 2 peptide epitope regions from the NS4 region. The
 CC antigen binds specifically with an antibody produced by a human infected
 CC by HCV. The peptide can detect a wide range of HCV infections with high
 CC sensitivity. (Updated on 27-AUG-2003 to correct OS field.)
 XX
 SQ Sequence 396 AA;

Query Match 5.9%; Score 7; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 50 PDKEVLY 56
 |||||
 Db 339 PDKEVLY 345

RESULT 864
 AAY80193
 ID AAY80193 standard; protein; 396 AA.
 XX
 AC AAY80193;
 XX
 DT 24-MAY-2000 (first entry)
 XX
 DE Hepatitis C virus chimeric antigen protein SEQ ID NO:11.
 XX
 KW Hepatitis C virus; HCV; antigen; antibody; infection; blood donation;
 KW screening; chimeric antigen.
 XX
 OS Hepatitis C virus.
 OS Synthetic.
 XX
 FN WO200007023-A1.
 XX
 PD 10-FEB-2000.
 XX
 PF 30-JUL-1999; 99WO-JP004129.
 XX
 PR 30-JUL-1998; 98JP-00216094.
 XX
 PA (ADLI-) ADVANCED LIFE SCI INST INC.

```
XX
PI Aoyagi K, Ohue C, Iida K, Yagi S;
XX WPI; 2000-195352/17.
DR N-PSDB; AAZ91497.
XX
XX Method of assaying for hepatitis C virus infection comprises binding
PT viral core antigen or anti-hepatitis C virus antibodies with a probe in
PT the presence of a surfactant.
XX
XX Example 11; Page 50-52; 55pp; Japanese.
XX
XX A method has been developed of assaying for hepatitis C virus (HCV). The
CC method comprises binding a viral core antigen or anti-HCV antibody to a
CC suitable probe in the presence of a non-ionic surfactant and/or a
CC cationic surfactant with secondary, tertiary or quaternary amine groups
CC and carbon chains with 10 or more carbon atoms. The assay is useful for
CC the detection of HCV infection in biological samples, especially in
CC donated blood. The method is highly sensitive for detecting HCV and
CC reduces the window period between negative and positive screening tests
CC for HCV on blood donated by a recently infected person. The present
CC sequence represents an HCV chimeric antigen protein given in an example
CC from the present invention
XX
XX Sequence 396 AA;
XX
Query Match 5.9%; Score 7; DB 3; Length 396;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 50 PDKEVLY 56
Db 339 PDKEVLY 345
|||||
XX
RESULT 865
ABO81715
ID ABO81715 standard; protein; 409 AA.
XX
AC ABO81715;
XX
XX 29-JUL-2004 (first entry)
XX
XX Pseudomonas aeruginosa polypeptide #13890.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX
XX N-PSDB; ABD15286.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 30461; 455pp; English.
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
```

```
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX Sequence 409 AA;
XX
Query Match 5.9%; Score 7; DB 7; Length 409;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 160 GGVLAAL 166
|||||
XX
RESULT 866
ABP80861
ID ABP80861 standard; protein; 411 AA.
XX
AC ABP80861;
XX
XX 07-MAR-2003 (first entry)
XX
XX N. gonorrhoeae amino acid sequence SEQ ID 8252.
XX
XX Antibacterial; infection; vaccine; gene therapy.
XX
XX Neisseria gonorrhoeae.
XX
XX WO200279243-A2.
XX
XX 10-OCT-2002.
XX
XX 12-FEB-2002; 2002WO-IB002069.
XX
XX 12-FEB-2001; 2001GB-00003424.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Fontana MR, Pizza M, Massignani V, Monaci E;
XX
XX WPI; 2003-058415/05.
XX
XX N-PSDB; ABZ41831.
XX
XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.
XX
XX Disclosure; Page 790; 815pp; English.
XX
XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.
CC Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention
XX
XX Sequence 411 AA;
XX
Query Match 5.9%; Score 7; DB 6; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 27
XX |||||
Db 79 LAALAA 85

RESULT 867
ABU49671
ID ABU49671 standard; protein; 414 AA.
XX AC
XX ABU49671;
XX 19-JUN-2003 (first entry)
XX DE
XX Protein encoded by Prokaryotic essential gene #35198.
DE Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX KW
XX Vibrio cholerae.
XX OS
XX WO200277183-A2.
XX PN
XX PD
XX 03-OCT-2002.
XX PF
XX 21-MAR-2002; 2002WO-US009107.
XX PR
XX 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX XX
XX (ELIT-) ELITRA PHARM INC.
XX PI
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zykkind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX WPI; 2003-029926/02.
XX DR N-PSDB; ACA53541.
XX PT
XX New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS
XX Claim 25; SEQ ID NO 77595; 1766pp; English.
XX CC
XX The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC compound's activity; (11) a culture comprising strains in which the gene
XX CC product is overexpressed or underexpressed; (12) determining the extent
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX XX
SQ Sequence 414 AA;
Query Match 5.9%; Score 7; DB 6; Length 414;
Best Local Similarity 100.0%; Pred. No. 5.4e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 VLLGGVL 21
XX |||||
Db 247 VLLGGVL 253

RESULT 868
ADS28776
ID ADS28776 standard; protein; 418 AA.
XX AC
XX ADS28776;
XX DT 02-DEC-2004 (first entry)
XX DE
XX Bacterial polypeptide #17809.
XX KW
XX Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.
XX OS
XX Bacteria.
XX PN US2003233675-A1.
XX XX
XX 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX PT
XX New recombinant DNA construct comprising a promoter positioned to provide
XX PT for expression of a polynucleotide encoding a polypeptide from a
XX PT microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 17809; 122pp; English.
XX CC
XX The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 418 AA;

Query Match 5.9%; Score 7; DB 8; Length 418;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
Db 133 VLALAA 139
|||||

RESULT 869

AAW37376
ID AAW37376 standard; protein; 421 AA.

XX AC AAW37376;

DT 27-AUG-2003 (revised)

DT 11-MAR-1998 (first entry)

XX Hepatitis C virus chimeric peptide antigen.

XX Hepatitis C virus; HCV; chimeric; antigen; detection; core region;
KW epitope; NS3; NS4; infection.

XX Synthetic.

OS Hepatitis C virus.

OS Chimeric.

XX JP09278794-A.

XX 28-OCT-1997.

XX 10-FEB-1997; 97JP-00027015.

XX 09-FEB-1996; 96JP-00024045.

XX (TOFU) TONEN CORP.

XX WPI; 1998-022248/03.

XX N-PSDB; AAT97237.

XX New chimaeric peptide antigen derived from hepatitis C virus protein -
PT useful for detecting HCV infections.

XX Claim 3; Page 19-21; 30pp; Japanese.

XX The present sequence represents a Hepatitis C virus (HCV) chimeric
CC peptide antigen which comprises at least 2 peptide epitope regions from
CC the HCV polypeptide core region, 2 peptide epitope regions from the NS3
CC region and at least 2 peptide epitope regions from the NS4 region. The
CC antigen binds specifically with an antibody produced by a human infected
CC by HCV. The peptide can detect a wide range of HCV infections with high
CC sensitivity. (Updated on 27-AUG-2003 to correct OS field.)

XX SQ Sequence 421 AA;

Query Match 5.9%; Score 7; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 361 PDKEVLY 367
|||||

RESULT 870

ABP78409

ID ABP78409 standard; protein; 421 AA.

XX AC ABP78409;

DT 07-MAR-2003 (first entry)

XX N. gonorrhoeae amino acid sequence SEQ ID 3348.

XX Antibacterial; infection; vaccine; gene therapy.

XX Neisseria gonorrhoeae.

XX WO200279243-A2.

XX 10-OCT-2002.

XX 12-FEB-2002; 2002WO-IB002069.

XX 12-FEB-2001; 2001GB-00003424.

XX (CHIR-) CHIRON SPA.

XX Fontana MR, Pizza M, Masignani V, Monaci E;

XX WPI; 2003-059415/05.

XX N-PSDB; ABZ39379.

XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a
PT medicament for treating or preventing N. gonorrhoeae infection.

XX Disclosure; Page 434; 815pp; English.

XX The present invention relates to proteins from Neisseria gonorrhoeae.
CC Also disclosed are the nucleic acid molecules encoding the proteins and
CC antibodies that specifically bind to the proteins. The composition
CC comprising the protein, nucleic acid or antibody is useful for the
CC manufacture of a medicament for treating or preventing N. gonorrhoeae
CC infection, this may be in the form of a vaccine or gene therapy.

XX Sequences given in records ABP76736-ABP81046 represent nucleic acid
CC molecules of the invention

XX SQ Sequence 421 AA;

Query Match 5.9%; Score 7; DB 6; Length 421;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47

Db 292 ELGGKPA 298
|||||

RESULT 871

ADI21224

ID ADI21224 standard; protein; 421 AA.

XX AC ADI21224;

DT 15-APR-2004 (first entry)

XX Novel human protein #199.

XX forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;

ligament tissue growth; nerve tissue growth; regeneration; wound healing; tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.

Homo sapiens.

WO2003025148-A2.

27-MAR-2003.

19-SEP-2002; 2002WO-US029964.

19-SEP-2001; 2001US-0323739P.

13-SEP-2002; 2002US-00323739.

(HYSE-) HYSEQ INC.

Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J; Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D; Haley-Vicente D;

WPI; 2003-354603/33.

N-PSDB; ADI21940.

New polynucleotides and secreted proteins, useful for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve tissue growth or regeneration, in wound healing, and in tissue repair and replacement.

Claim 20; SEQ ID NO 475; 156pp; English.

The invention relates to an isolated polynucleotide encoding a polypeptide with biological activity. The polynucleotides and polypeptides are useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders and other traits, to assess biodiversity, as nutritional sources or supplements. The polynucleotides may also be used as molecular weight markers, chromosome markers or map related gene positions, or as an antigen to raise anti-DNA antibodies or elicit immune response. The polypeptides are useful for raising antibodies, as markers for tissues in which the corresponding polypeptide is expressed, for re-engineering damaged or diseased tissues, for treating myeloid or lymphoid cell disorders, in bone cartilage, tendon, ligament and/or nerve tissue growth or regeneration, in wound healing, in tissue repair and replacement, in healing of burns, incisions and ulcers, and in treating cancer. The present sequence represents the amino acid sequence of a novel human protein.

Sequence 421 AA;

Query Match 5.9%; Score 7; DB 7; Length 421; Best Local Similarity 100.0%; Pred. No. 5.5e+02; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
|||||
Db 410 GGVLAAL 416

RESULT 872
AAU27812
ID AAU27812 standard; protein; 426 AA.
XX
AC AAU27812;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human full-length polypeptide sequence #137.
XX
KW Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat; mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia; cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation; nervous system disorder; inflammatory disorder; cell differentiation; angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;

Genetic disorder; bone regeneration; tendon; ligament; tissue repair; cytostatic; antirheumatic; antiarthritic; vulnary; antinflammatory; antibacterial; immunosuppressive; vasotropic; antiparkinsonian; neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic; immunostimulant; analgesic; gene therapy.

Homo sapiens.

WO200164834-A2.

07-SEP-2001.

26-FEB-2001; 2001WO-US004926.

28-FEB-2000; 2000US-00515126.

18-MAY-2000; 2000US-00577409.

17-JUN-2000; 2000US-00597707.

14-JUL-2000; 2000US-00616807.

19-SEP-2000; 2000US-00664641.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F; Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C; Drmanac R;

WPI; 2001-589862/66.

N-PSDB; AAS44712.

Novel polypeptides and nucleic acids obtained from cDNA libraries prepared from various human tissues, for diagnosis, treatment of cancer, neurological, inflammatory disorders and for use in arrays for detection.

Claim 10; SEQ ID NO 309; 153pp; English.

Sequences AAU27676-AAU28019 represent full-length polypeptides and contig polypeptides of the invention. The proteins and their associated DNA sequences are useful for the treatment, diagnosis and prevention of various types of disorder in a mammalian subject such as a human, dog, monkey, mouse, hamster or rat. The disorders include cancers such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as multiple sclerosis, connective tissue disease, rheumatoid arthritis, diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system disorders such as Parkinson's disease, Alzheimer's disease, Huntington's chorea, anyotrophic lateral sclerosis, spinal muscular atrophy and Wernicke disease, inflammatory disorders such as nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory bowel disease. The sequences exhibit activity relating to angiogenesis, cell proliferation, cell differentiation, stem cell growth factor, activin or inhibin. Therefore, they can be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, accidental damage or genetic disorders. The sequences may also be used for regeneration of bone, cartilage, tendons and ligaments and in tissue repair and burn healing. Note: Some sequences for this patent did not form part of the printed specification, but were obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 426 AA;

Query Match 5.9%; Score 7; DB 4; Length 426; Best Local Similarity 100.0%; Pred. No. 5.5e+02; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 PAIVPDK 52
|||||
Db 212 PAIVPDK 218

RESULT 873
AAU42357
ID AAU42357 standard; protein; 426 AA.
XX

RESULT 875

ABU28800
ID ABU28800 standard; protein; 428 AA.

XX AC ABU28800;
XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #14327.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Escherichia coli.

XX PN WO200277183-A2.
XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.

XX PR (ELIT-) ELITRA PHARM INC.
XX PA

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA32670.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 56724; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway;
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC product is overexpressed or underexpressed; (11) a culture comprising strains in which the gene
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 428 AA;

Query Match

Best Local Similarity 5.9%; Score 7; DB 6; Length 428;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QOAVIEP 102
Db 259 QOAVIEP 265

RESULT 876

ABU37936
ID ABU37936 standard; protein; 431 AA.

XX AC ABU37936;
XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #23463.
XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX OS Neisseria meningitidis.

XX PN WO200277183-A2.
XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.
XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.

XX PR (ELIT-) ELITRA PHARM INC.
XX PA

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX DR WPI; 2003-029926/02.
XX DR N-PSDB; ACA41806.

XX PT New antisense nucleic acids, useful for identifying proteins or screening
XX PT for homologous nucleic acids required for cellular proliferation to
XX PT isolate candidate molecules for rational drug discovery programs.
XX PS Claim 25; SEQ ID NO 65860; 1766pp; English.
XX CC The invention relates to an isolated nucleic acid comprising any one of
XX CC the 6213 antisense sequences given in the specification where expression
XX CC of the nucleic acid inhibits proliferation of a cell. Also included are:
XX CC (1) a vector comprising a promoter operably linked to the nucleic acid
XX CC encoding a polypeptide whose expression is inhibited by the antisense
XX CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX CC polypeptide or its fragment whose expression is inhibited by the
XX CC antisense nucleic acid; (4) an antibody capable of specifically binding
XX CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX CC proliferation or the activity of a gene in an operon required for
XX CC proliferation; (7) identifying a compound that influences the activity of
XX CC the gene product or that has an activity against a biological pathway;
XX CC required for proliferation, or that inhibits cellular proliferation; (8)
XX CC identifying a gene required for cellular proliferation or the biological
XX CC pathway in which a proliferation-required gene or its gene product lies
XX CC or a gene on which the test compound that inhibits proliferation of an
XX CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX CC product is overexpressed or underexpressed; (11) a culture comprising strains in which the gene
XX CC to which each of the strains is present in a culture or collection of
XX CC strains; or (13) identifying the target of a compound that inhibits the
XX CC proliferation of an organism. The antisense nucleic acids are useful for
XX CC identifying proteins or screening for homologous nucleic acids required
XX CC for cellular proliferation to isolate candidate molecules for rational
XX CC drug discovery programs, or for screening homologous nucleic acids
XX CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX CC the target prokaryotic essential genes. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences

CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 431 AA;

Query Match 5.9%; Score 7; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 41 ELGGKPA 47
Db 302 ELGGKPA 308
|||||||

RESULT 877
ID ABU37325 standard; protein; 431 AA.
XX
AC ABU37325;
XX
DT 23-OCT-2003 (revised)
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by Prokaryotic essential gene #22852.
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX *Neisseria gonorrhoeae*.
XX
XX WO200277183-A2.
XX
PD 03-OCT-2002.
XX
PF 21-MAR-2002; 2002WO-US009107.
XX
PR 21-MAR-2001; 2001US-00815242.
PR 06-SEP-2001; 2001US-00948993.
PR 25-OCT-2001; 2001US-0342923P.
PR 08-FEB-2002; 2002US-00072851.
PR 06-MAR-2002; 2002US-0362699P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA41195.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 65249; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing a polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 431 AA;

Query Match 5.9%; Score 7; DB 6; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 41 ELGGKPA 47
Db 302 ELGGKPA 308
|||||||

RESULT 878
ID ADP08253 standard; protein; 431 AA.
XX
AC ADP08253;
XX
DT 26-AUG-2004 (first entry)
XX
DE *Neisseria meningitidis* MC58 OMV-related membrane protein - SEQ ID 86.
XX
KW outer-membrane vesicle; antibacterial; antiinflammatory;
KW meningococcal protein trafficking; localisation; infection; vaccine;
KW gene therapy.
XX
OS *Neisseria meningitidis* MC58.
XX
PN WO2004046177-A2.
XX
PD 03-JUN-2004.
XX
PF 17-NOV-2003; 2003WO-IB006281.
XX
PR 15-NOV-2002; 2002GB-00026734.
PR 27-MAR-2003; 2003GB-00007131.
XX
PA (CHIR) CHIRON SRL.
XX
PI Norais N, Grandi G;
XX
DR WPI; 2004-420615/39.
XX
PT New compositions having outer-membrane vesicles and proteins from
PT *Neisseria meningitidis*, useful in the field of meningococcal
PT biochemistry, in particular for preventing and/or treating meningococcal
PT infections.
XX
PS Claim 9; SEQ ID NO 86; 79pp; English.
XX
XX The invention relates to a novel composition comprising outer-membrane
CC vesicles (OMV) prepared from a first strain of *Neisseria meningitidis* and
CC 1 or more proteins which are present in OMVs prepared from a second

CC strain of *N. meningitidis*, but which are not present in OMVs prepared
 CC from the first strain. The composition of the invention demonstrates
 CC antibacterial and anti-inflammatory activities and may be useful in the
 CC field of meningococcal biochemistry, in particular the trafficking and
 CC localisation of meningococcal proteins, as well as in the prevention or
 CC treatment of meningococcal infections, possibly via the production of a
 CC vaccine or gene therapy. The current sequence is that of a *Neisseria*
 CC meningitidis MC58 outer-membrane vesicle (OMV)-related membrane protein
 CC of the invention.
 XX
 XX SQ Sequence 431 AA;

Query Match 5.9%; Score 7; DB 8; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ELGGKPA 47
 Db 302 ELGGKPA 308
 |||||

RESULT 879
 ADA36359
 ID ADA36359 standard; protein; 433 AA.

XX
 AC ADA36359;

XX
 DT 20-NOV-2003 (first entry)

XX
 DE Acinetobacter baumannii protein #3520.

XX
 KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 XX plant biocontrol agent.

XX
 OS Acinetobacter baumannii.

XX
 PN US562958-B1.

XX
 PD 13-MAY-2003.

XX
 PF 04-JUN-1999; 99US-00328352.

XX
 PR 09-JUN-1998; 98US-008701P.

XX
 PA (GENO-) GENOME THERAPEUTICS CORP.

XX
 PI Breton G, Bush D;

XX
 DR WPI; 2003-576092/54.

XX
 DR N-PSDB; ADA32233.

XX
 PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 PT for diagnosing a bacterial disease, as components of antibacterial
 PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 PT plants.

XX
 PS Example; SEQ ID NO 7646; 328pp; English.

XX
 CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
 CC The A. baumannii nucleic acids and polypeptides are useful as reagents
 CC for diagnosing a bacterial disease, as components of antibacterial
 CC vaccines, as targets for antibacterial drugs, to detect the presence of
 CC A. baumannii and other Acinetobacter species in a sample, in screening
 CC compounds for the ability to interfere with the A. baumannii life cycle
 CC or to inhibit A. baumannii infection, and as biocontrol agents for
 CC plants. The present sequence represents the amino acid sequence of an A.
 CC baumannii protein.
 XX
 XX SQ Sequence 433 AA;

Query Match 5.9%; Score 7; DB 6; Length 433;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LYQQYDE 61
 Db 103 LYQQYDE 109
 |||||

RESULT 880
 ABO61746
 ID ABO61746 standard; protein; 438 AA.

XX
 AC ABO61746;

XX
 DT 29-JUL-2004 (first entry)

XX
 DE Klebsiella pneumoniae polypeptide seqid 8263.

XX
 KW Recombinant expression vector; transcription regulatory element;
 XX Klebsiella pneumoniae protein; antibacterial; vaccine.

XX
 OS Klebsiella pneumoniae.

XX
 PN US6610836-B1.

XX
 PD 26-AUG-2003.

XX
 PF 27-JAN-2000; 2000US-00489039.

XX
 PR 29-JAN-1999; 99US-0117747P.

XX
 PA (GENO-) GENOME THERAPEUTICS CORP.

XX
 PI Breton GL, Osborne M;

XX
 DR WPI; 2003-895346/82.

XX
 DR N-PSDB; ACH95297.

XX
 PT New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 XX preparing a vaccine composition against Klebsiella pneumoniae.

XX
 PS Disclosure; SEQ ID NO 8263; 932pp; English.

XX
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 XX SQ Sequence 438 AA;

Query Match 5.9%; Score 7; DB 7; Length 438;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
 Db 212 GGVLAAL 218
 |||||

RESULT 881
 ADS30300
 ID ADS30300 standard; protein; 441 AA.

XX
 AC ADS30300;

XX
 DT 02-DEC-2004 (first entry)

XX
 DE Bacterial polypeptide #19333.

XX
 KW Recombinant DNA construct; transformed plant; improved plant property;
 XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;

KW call cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 19333; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition. Improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 441 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 441;
Best Local Similarity 100.0%; Pred. No. 5.7e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
Oy 4 SADLEVT 10
Db 215 SADLEVT 221
RESULT 882
ADS42560
ID ADS42560 standard; protein; 452 AA.
XX
XX ADS42560;
XX

DT 02-DEC-2004 (first entry)
XX Bacterial polypeptide #20990.
XX Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX Bacteria.
XX US2003233675-A1.
XX 18-DEC-2003.
XX 20-FEB-2003; 2003US-00369493.
XX 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
XX (HINK/) HINKLE G J.
XX (SLAT/) SLATER S C.
XX (CHEN/) CHEN X.
XX (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 20990; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition. Improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 452 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 452;
Best Local Similarity 100.0%; Pred. No. 5.8e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
Oy 20 VLAALAA 26
Db 39 VLAALAA 45

RESULT 883
ABU49575
ID ABU49575 standard; protein; 453 AA.
XX AC
XX ABU49575;
XX
XX 19-JUN-2003 (first entry)
XX
XX Protein encoded by Prokaryotic essential gene #35102.
XX DE
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX KW
XX Vibrio cholerae.
XX OS
XX WO200277183-A2.
XX FN
XX
XX 03-OCT-2002.
XX PD
XX 21-MAR-2002; 2002WO-US009107.
XX PF
XX 21-MAR-2001; 2001US-00815242.
XX PR 06-SEP-2001; 2001US-00948993.
XX PR 25-OCT-2001; 2001US-0342923P.
XX PR 08-FEB-2002; 2002US-00072851.
XX PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
XX PA
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;
XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX DR N-PSDB; ACA53445.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX PT
XX
XX Claim 25; SEQ ID NO 77499; 1766pp; English.
XX PS
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 453 AA;

Query Match 5.9%; Score 7; DB 6; Length 453;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 296 VLAALAA 302

RESULT 884

ABU22514
ID ABU22514 standard; protein; 457 AA.

XX AC

XX ABU22514;

XX DT 19-JUN-2003 (first entry)

XX DE Protein encoded by Prokaryotic essential gene #8041.

XX KW Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX OS Burkholderia mallei.

XX XX

XX FN WO200277183-A2.

XX PD 03-OCT-2002.

XX PF 21-MAR-2002; 2002WO-US009107.

XX PR 21-MAR-2001; 2001US-00815242.

XX PR 06-SEP-2001; 2001US-00948993.

XX PR 25-OCT-2001; 2001US-0342923P.

XX PR 08-FEB-2002; 2002US-00072851.

XX PR 06-MAR-2002; 2002US-0362699P.

XX XX

XX (ELIT-) ELITRA PHARM INC.

XX PA

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zvekind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX
XX WPI; 2003-029926/02.
XX DR N-PSDB; ACA26384.XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX PT
XX
XX Claim 25; SEQ ID NO 50438; 1766pp; English.

XX PS

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

XX of the nucleic acid inhibits proliferation of a cell. Also included are:

XX (1) a vector comprising a promoter operably linked to the nucleic acid

XX encoding a polypeptide whose expression is inhibited by the antisense

XX nucleic acid; (2) a host cell containing the vector; (3) an isolated

XX antisense nucleic acid; (4) an antibody capable of specifically binding

XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

XX proliferation or the activity of a gene in an operon required for

XX proliferation; (7) identifying a compound that influences the activity of

XX the gene product or that has an activity against a biological pathway

XX required for proliferation, or that inhibits cellular proliferation; (8)

XX identifying a gene required for cellular proliferation or the biological

XX pathway in which a proliferation-required gene or its gene product lies

XX or a gene on which the test compound that inhibits proliferation of an

XX organism acts; (9) manufacturing an antibiotic; (10) profiling a

XX compound's activity; (11) a culture comprising strains in which the gene

XX product is overexpressed or underexpressed; (12) determining the extent

XX to which each of the strains is present in a culture or collection of

XX strains; or (13) identifying the target of a compound that inhibits the

XX proliferation of an organism. The antisense nucleic acids are useful for

XX identifying proteins or screening for homologous nucleic acids required

XX for cellular proliferation to isolate candidate molecules for rational

XX drug discovery programs, or for screening homologous nucleic acids

XX required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
XX *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
XX the target prokaryotic essential genes. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences

CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 457 AA;

Query Match 5.9%; Score 7; DB 6; Length 457;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 154 VLLGGVL 160
|||||

RESULT 885
ABO73936
ID ABO73936 standard; protein; 457 AA.

AC ABO73936;

XX 29-JUL-2004 (first entry)

XX Pseudomonas aeruginosa polypeptide #6111.

XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.

XX Pseudomonas aeruginosa.

XX US6551795-B1.

XX 22-APR-2003.

XX 18-FEB-1999; 99US-00252991.

XX 18-FEB-1998; 98US-0074788P.

XX 27-JUL-1998; 98US-0094190P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Rubenfield MJ, Nollong J, Deloughery C, Bush D;

XX WPI; 2003-615309/58.

XX N-PSDB; ABD07507.

XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,

XX useful as molecular targets for diagnostics, prophylaxis and treatment of

XX pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 22682; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the
XX polynucleotides encoding them. The sequences are useful in diagnosis and
XX therapy of pathological conditions, as molecular targets for diagnostics,
XX prophylaxis and treatment of pathological conditions resulting from a
XX bacterial infection, for evaluating a compound, such as a polypeptide,
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
XX effective antibacterial targets, as targets for antibacterial drugs,
XX including anti-P. aeruginosa drugs, as templates for recombinant
XX production of P. aeruginosa-derived peptides or polypeptides, as target
XX components for diagnosis and/or treatment of P. aeruginosa-caused
XX infection, and in detection of P. aeruginosa sequences or other sequences
XX of Pseudomonas species using biochip technology. Sequences ABO7826-
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format from USPTO at
XX seqdata.uspto.gov/sequence.html

XX SQ Sequence 457 AA;

Query Match 5.9%; Score 7; DB 7; Length 457;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGLLQR 92
Db 142 VLGLLQR 148
|||||

RESULT 886
ADX96169
ID ADX96169 standard; protein; 458 AA.

XX AC ADX96169;

XX 21-APR-2005 (first entry)

XX Plant full length insert polypeptide seqid 58833.

XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content.

XX Unidentified.

XX US2004034888-A1.

XX 19-FEB-2004.

XX 28-APR-2003; 2003US-00425114.

XX 06-MAY-1999; 99US-00304517.

XX 05-NOV-2001; 2001US-00985678.

XX (LIUJ/) LIU J.

XX (ZHOU/) ZHOU Y.

XX (KOA/) KOVALIC D K.

XX (SCRE/) SCREEN S E.

XX (TAB/) TABASKA J E.

XX (CAOY/) CAO Y.

XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;

XX WPI; 2004-180133/17.

XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.

XX Claim 1; SEQ ID NO 58833; 15pp; English.

XX The invention describes a recombinant DNA construct comprising a
XX polynucleotide consisting of a sequence encoding an amino acid sequence
XX available in electronic form from the US patent office at
XX ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide
XX of the invention are also useful in physical arrays of molecules and as
XX plant breeding markers. The recombinant DNA construct is useful for
XX improving plant tolerance to cold, heat, drought, herbicides, extreme
XX osmotic conditions, pathogens or pests, for manipulating growth rate in
XX plant cells by modification of the cell cycle pathway, for conferring
XX increased resistance to plant disease, for producing galactomannan,
XX lignin or plant growth regulators, for increasing the rate of homologous
XX recombination in plants, for improving yield by modification of
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake

CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This is the amino acid sequence of a plant full length insert
CC polypeptide that can be used in the recombinant DNA construct of the
CC invention.

XX
SQ Sequence 458 AA;
Query Match 5.9%; Score 7; DB 8; Length 458;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLALAA 26
Db 85 VLALAA 91
|||||

RESULT 887
ADJ53182
ID ADJ53182 standard; protein; 473 AA.
XX
AC ADJ53182;
XX
DT 06-MAY-2004 (first entry)
XX
DE Rebeccamycin biosynthetic membrane transport protein OrfR1.
XX
KW Rebeccamycin biosynthetic gene cluster; rebeccamycin production;
KW indolocarbazole production; ATCC39243; Streptomyces albus;
KW antitumor agent; cytostatic; Gram positive bacterium; antibacterial;
KW antibiotic; OrfR1; membrane transport protein.
XX

OS Lechevalieria aerocolonigenes; ATCC39243.

XX
XX WO200303706-A1.

PN
XX 24-APR-2003.

PD
XX 17-OCT-2002; 2002WO-ES000492.

PF
XX 19-OCT-2001; 2001ES-00002312.

PR
XX (UYOV-) UNIV OVIEDO.

PA Sanchez Reillo C, Fernandez Brana A, Salas Fernandez JA;
PI Mendez Fernandez C;

XX
XX WPI; 2003-393533/37.

DR
XX N-PSDB; ADJ53165.

XX
XX Production of indolo-carbazole antitumor agents, especially rebeccamycin
or derivatives, by culturing host cells transformed with vector
containing DNA from Saccharothrix aerocolonigenes.

PS Claim 10; SEQ ID NO 18; 76pp; Spanish.

XX
XX The invention relates to a method for the production of indolocarbazoles
in Streptomyces (especially Streptomyces albus) using rebeccamycin
biosynthetic genes isolated from Saccharothrix aerocolonigenes ATCC39243.
CC The invention also relates to the Saccharothrix aerocolonigenes ATCC39243
rebeccamycin biosynthetic gene cluster (ADJ53165) and sequences at least
80% homologous to it; nucleic acids encoding one or more rebeccamycin
biosynthetic proteins; the 18 rebeccamycin biosynthetic proteins
(ADJ53166-ADJ53183) encoded by the gene cluster; and vectors and host
cells comprising nucleic acid sequences of the invention.
CC Indolocarbazoles (especially rebeccamycin) and their derivatives and
precursors are useful as antitumor agents. Rebeccamycin also shows
antibacterial activity against Gram positive bacteria such as
Staphylococcus aureus, Micrococcus luteus and Streptococcus faecalis. The
recombinant Streptomyces albus of the invention produce increased yields
of indolocarbazoles compared with Saccharothrix aerocolonigenes
ATCC39243, and may permit the production of novel indolocarbazole
compounds (e.g., halogenated derivatives). The present sequence

CC represents a specifically claimed rebeccamycin biosynthetic protein from
CC Saccharothrix aerocolonigenes ATCC39243.
XX
SQ Sequence 473 AA;

Query Match 5.9%; Score 7; DB 7; Length 473;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
Db 153 VLLGGVL 159
|||||

RESULT 888
ABM92710
ID ABM92710 standard; protein; 477 AA.

XX
AC ABM92710;

XX
DT 02-JUN-2005 (first entry)

XX
DE M. xanthus protein sequence, seq id 11909.

XX
KW Transgenic plant; DNA replication; gene regulation; gene expression.

OS Myxococcus xanthus.

XX
XX US6833447-B1.

PN
XX 21-DEC-2004.

PD
XX 10-JUL-2001; 2001US-00902540.

PF
XX 10-JUL-2000; 2000US-0217883P.

PR
XX (MONS) MONSANTO TECHNOLOGY LLC.

PA Goldman BS, Hinkle GJ, Slater SC, Wiegand RC;

XX
XX WPI; 2005-028716/03.

DR
XX New substantially purified Myxococcus xanthus nucleic acid molecule
encoding a nitrite reductase, useful for determining gene expression,
identifying mutations in a gene of interest, and for constructing
mutations in a gene of interest.

PS Example 2; SEQ ID NO 11909; 25pp; English.

XX
XX The invention relates to a substantially purified nucleic acid molecule
encoding a nitrite reductase of SEQ ID NO. 11926. Further disclosed is a
recombinant DNA construct for expression of a nitrite reductase gene in a
plant cell, and a plant cell comprising the recombinant DNA construct.
CC The nucleic acid is useful for determining gene expression, identifying
mutations in a gene of interest, and for constructing mutations in a gene
of interest. Sequences given in records for SEQ IDs 9692-16825 represent
a group of 7134 Myxococcus xanthus proteins and peptides. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from USPTO

SQ Sequence 477 AA;

Query Match 5.9%; Score 7; DB 9; Length 477;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLALAA 26
Db 465 VLALAA 471
|||||

RESULT 889
ADT89472

XX The present invention relates to a nematode phosphoglycerate mutase (PGM)
 CC and its encoding nucleic acid molecule. The nucleic acid molecule of the
 CC invention is useful in preparing a composition for combating diseases or
 CC infestations caused by nematodes. The invention is also useful in gene
 CC therapy. The present sequence is a Meloidogyne incognita phosphoglycerate
 CC mutase (PGM) like protein. Note: This sequence is defined as
 CC Caenorhabditis elegans phosphoglycerate mutase (PGM) in the column 11 of
 CC the specification.
 XX Sequence 526 AA;

Query Match 5.9%; Score 7; DB 8; Length 526;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 NWQKLEA 113
 DB 45 NWQKLEA 51
 |||||

RESULT 892

ADU48511
 ID ADU48511 standard; protein; 526 AA.

XX AC ADU48511;

XX DT 27-JAN-2005 (first entry)

XX DE Meloidogyne incognita phosphoglycerate mutase (PGM) like protein #1.

XX KW Nematode; phosphoglycerate mutase; PGM; combating disease; infestation;
 XX gene therapy; nematocide; enzyme.

XX OS Meloidogyne incognita.

XX FH Key Location/Qualifiers

XX FT Domain 3..75

XX FT /note = PGM-like protein domain

XX FT Domain 83..306

XX FT /note = PGM-like protein domain

XX FT Domain 317..523

XX FT /note = PGM-like protein domain

XX US6818433-B1.

XX PN 16-NOV-2004.

XX PD 26-FEB-2002; 2002US-00082894.

XX PF 27-FEB-2001; 2001US-0271781P.

XX PR (DIVE-) DIVERGENCE INC.

XX PA Klook AP, Williams DJ, Salmon B, Bradley JD;

XX PI WPI; 2004-793565/78.

XX DR N-PSDB; ADU48510.

XX FT New phosphoglycerate mutase nucleic acid useful in preparing a
 FT composition for combating diseases or infestations caused by nematodes.

XX PS Claim 1; SEQ ID NO 2; 29pp; English.

XX The present invention relates to a nematode phosphoglycerate mutase (PGM)
 CC and its encoding nucleic acid molecule. The nucleic acid molecule of the
 CC invention is useful in preparing a composition for combating diseases or
 CC infestations caused by nematodes. The invention is also useful in gene
 CC therapy. The present sequence is a Meloidogyne incognita phosphoglycerate
 CC mutase (PGM) like protein.

XX Sequence 526 AA;

Query Match 5.9%; Score 7; DB 8; Length 526;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 107 NWQKLEA 113
 DB 45 NWQKLEA 51
 |||||

RESULT 893

ABM67618
 ID ABM67618 standard; protein; 530 AA.

XX AC ABM67618;

XX DT 20-NOV-2003 (first entry)

XX DE Photorhabdus luminescens protein sequence #715.

XX KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.

XX OS Photorhabdus luminescens.

XX PN WO200294867-A2.

XX PD 28-NOV-2002.

XX PF 07-FEB-2002; 2002WO-IB003040.

XX PR 07-FEB-2001; 2001FR-00001659.

XX PA (INSP) INST PASTEUR.

XX PA (CNRS) CNRS CENT NAT RECH SCI.

XX PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;

XX DR WPI; 2003-148459/14.

XX PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.

XX PS Claim 2; SEQ ID NO 715; 1205pp; French.

XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins

XX Sequence 530 AA;

Query Match 5.9%; Score 7; DB 6; Length 530;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 246 GGVLAAL 252
|||||

RESULT 894
AAU27984
ID AAU27984 standard; protein; 550 AA.
XX
AC AAU27984;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human contig polypeptide sequence #137.
XX
XX Mammal; human; rhesus monkey; baker's yeast; fission yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; lymphoma; neuroblastoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation;
KW angiogenesis; stem cell growth factor; activin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytotatic; antirheumatic; antiarthritic; vulnary; antiinflammatory;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW neuroprotective; osteopathic; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX WO200164834-A2.
PN
XX
XX 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US004926.
XX
XX 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
PR 17-JUN-2000; 2000US-00597707.
PR 14-JUL-2000; 2000US-00616807.
PR 19-SEP-2000; 2000US-00666461.
XX
XX (HYSE-) HYSEQ INC.
FA
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Wang J, Ma Y, Wang D, Chen R, Xu C;
PI Drmanac R;
XX
XX WPI; 2001-599862/66.
DR N-PSDB; RAAS44884.
XX
XX Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of cancer,
PT neurological, inflammatory disorders and for use in arrays for detection.
XX
XX Claim 10; Page 142-143; 153pp; English.
PS
XX Sequences AAU27676-AAU28019 represent full-length polypeptides and contig
CC polypeptides of the invention. The proteins and their associated DNA
CC sequences are useful for the treatment, diagnosis and prevention of
CC various types of disorder in a mammalian subject such as a human, dog,
CC monkey, mouse, hamster or rat. The disorders include cancers such as
CC leukaemia, lymphoma and neuroblastoma, autoimmune disorders such as
CC multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC activin or inhibin. Therefore, they can be used to manipulate stem cells

CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 550 AA;
Query Match 5.9%; Score 7; DB 4; Length 550;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 46 PAIVPDK 52
Db 336 PAIVPDK 342
|||||

RESULT 895
ADZ51375
ID ADZ51375 standard; protein; 553 AA.
XX
AC ADZ51375;
XX
DT 30-JUN-2005 (first entry)
XX
DE Amino acid sequence of ovarian cancer marker M715.
XX
KW cytostatic; gene therapy; ovarian cancer; ovarian cancer marker; M715;
KW solute carrier family 39 protein.
XX
OS Homo sapiens.
XX
XX WO2005034732-A2.
PN
XX 21-APR-2005.
PD
XX 07-OCT-2004; 2004WO-US033166.
PR
XX 07-OCT-2003; 2003US-0509171P.
PR
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Endege WO, Ford D, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S;
PI Monahan JB, Schlegel R, Xu YY, Zhao X;
XX
XX WPI; 2005-306219/31.
DR N-PSDB; ADZ51374.
XX
XX Assessing whether a patient is afflicted with ovarian cancer comprises
PT determining a significant difference between the levels of expression of
PT an ovarian cancer marker in the patient sample and the sample from a
PT control subject.
XX
XX Claim 36; SEQ ID NO 32; 164pp; English.
PS
XX The specification describes a method of assessing whether a patient is
CC afflicted with ovarian cancer. The method comprises determining the
CC presence of a significant difference between the levels of expression of
CC an ovarian cancer marker in the patient sample and the sample from a
CC control subject. The method of the invention is useful for assessing,
CC diagnosing, preventing or treating ovarian cancer. The markers may also
CC be used in screening for agents that may treat or prevent ovarian cancer.
CC The present sequence represents the marker M715, which is solute carrier
CC family 39 protein variant 2.
XX
SQ Sequence 553 AA;
Query Match 5.9%; Score 7; DB 9; Length 553;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
DB 86 GGVLAAL 92

RESULT 896

AAB10080
ID AAB10080 standard; protein; 563 AA.

XX AC AAB10080;

XX DT 03-NOV-2000 (first entry)

XX DE F. bidentis glutamate/malate translocator protein.

XX KW Glutamate/malate translocator; herbicide; transamination; glutamate;
KW dicarboxylic acid transport inhibitor; tissue specific expression; plant.

XX OS Flaveria bidentis.

XX PN WO200031281-A2.

XX PD 02-JUN-2000.

XX PF 22-NOV-1999; 99WO-EP008960.

XX PR 21-NOV-1998; 98DE-01053778.

XX PA (BADI) BASF AG.

XX PI Fluegge U, Weber A, Westhoff P, Dressen U;

XX DR WPI; 2000-411957/35.

XX DR N-PSDB; AAA40318.

XX PT New DNA encoding glutamate-malate transporter, useful for producing
PT transgenic plants with altered nitrogen metabolism, particularly
PT increased protein content.

XX PS Disclosure; Page 33-35; 41pp; German.

XX CC This invention describes novel DNA sequences (I) containing the coding
CC region of a glutamate-malate transporter (II). The products of the
CC invention have herbicidal activity. (II) exports glutamate, which is the
CC main amino donor for a wide variety of transamination reactions, from
CC chloroplasts and is a key regulator in intracellular partitioning of
CC amino group donors and acceptors. (I) and its derivatives are used: (i)
CC to express (II) in prokaryotic or eukaryotic cells; (ii) for isolation of
CC DNA encoding (II); (iii) for combination with target sequences to provide
CC expression of (II) in other cell compartments or membrane systems; (iv)
CC to identify agents that inhibit transport of dicarboxylic acids across
CC the inner plastid membrane (potential total herbicides); (v) to isolate
CC corresponding genomic clones; (vi) to produce sequences that encode
CC dicarboxylic acid translocators with altered substrate specificity or to
CC impart resistance to inhibitors of (iv), for development of selective
CC herbicides; and (vii) to identify insertional mutants, for homologous
CC recombination and for expression of non-translatable RNA (by antisense,
CC co-suppression or ribozyme effects) to suppress one or more endogenous
CC plastid (II) in cells. The promoter region of (I) may be used to provide
CC tissue-specific expression of genes. Expression/suppression of (II)
CC modifies formation and transport of carbon skeletons for nitrogen
CC fixation, and thus transport of assimilable nitrogen, e.g. it may produce
CC plants with reduced carbohydrate content but increased content of organic
CC nitrogen (proteins, amino acids, alkaloids), particularly food or fodder
CC plants with increased protein content. This sequence represents the
CC Flaveria bidentis glutamate/malate translocator protein which is
CC described in the invention

XX SQ Sequence 563 AA;

Query Match 5.9%; Score 7; DB 3; Length 563;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 25
DB 495 GVLAAAL 501

RESULT 897

AAB10081
ID AAB10081 standard; protein; 564 AA.

XX AC AAB10081;

XX DT 03-NOV-2000 (first entry)

XX DE S. oleracea glutamate/malate translocator protein.

XX KW Glutamate/malate translocator; herbicide; transamination; glutamate;
KW dicarboxylic acid transport inhibitor; tissue specific expression; plant;
KW spinach.

XX OS Spinacia oleracea.

XX PN WO200031281-A2.

XX PD 02-JUN-2000.

XX PF 22-NOV-1999; 99WO-EP008960.

XX PR 21-NOV-1998; 98DE-01053778.

XX PA (BADI) BASF AG.

XX PI Fluegge U, Weber A, Westhoff P, Dressen U;

XX DR WPI; 2000-411957/35.

XX DR N-PSDB; AAA40319.

XX PT New DNA encoding glutamate-malate transporter, useful for producing
PT transgenic plants with altered nitrogen metabolism, particularly
PT increased protein content.

XX PS Disclosure; Page 39-41; 41pp; German.

XX CC This invention describes novel DNA sequences (I) containing the coding
CC region of a glutamate-malate transporter (II). The products of the
CC invention have herbicidal activity. (II) exports glutamate, which is the
CC main amino donor for a wide variety of transamination reactions, from
CC chloroplasts and is a key regulator in intracellular partitioning of
CC amino group donors and acceptors. (I) and its derivatives are used: (i)
CC to express (II) in prokaryotic or eukaryotic cells; (ii) for isolation of
CC DNA encoding (II); (iii) for combination with target sequences to provide
CC expression of (II) in other cell compartments or membrane systems; (iv)
CC to identify agents that inhibit transport of dicarboxylic acids across
CC the inner plastid membrane (potential total herbicides); (v) to isolate
CC corresponding genomic clones; (vi) to produce sequences that encode
CC dicarboxylic acid translocators with altered substrate specificity or to
CC impart resistance to inhibitors of (iv), for development of selective
CC herbicides; and (vii) to identify insertional mutants, for homologous
CC recombination and for expression of non-translatable RNA (by antisense,
CC co-suppression or ribozyme effects) to suppress one or more endogenous
CC plastid (II) in cells. The promoter region of (I) may be used to provide
CC tissue-specific expression of genes. Expression/suppression of (II)
CC modifies formation and transport of carbon skeletons for nitrogen
CC fixation, and thus transport of assimilable nitrogen, e.g. it may produce
CC plants with reduced carbohydrate content but increased content of organic
CC nitrogen (proteins, amino acids, alkaloids), particularly food or fodder
CC plants with increased protein content. This sequence represents the
CC spinach (Spinacia oleracea) glutamate/malate translocator protein which
CC is described in the invention

XX SQ Sequence 564 AA;

Query Match 5.9%; Score 7; DB 3; Length 564;

Best Local Similarity 100.0%; Pred. No. 7e+02;					
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	19	GVLAAALA 25			
Db	496	GVLAAALA 502			
 RESULT 898					
ABB65162					
ID	ABB65162 standard; protein; 567 AA.				
XX	AC	AC			
XX	AB	B65162;			
XX	26-MAR-2002	(first entry)			
XX	DT	DT			
XX	XX	XX			
DE	Drosophila melanogaster polypeptide SEQ ID NO 22278.				
XX	XX	XX			
KW	Drosophila; developmental biology; cell signalling; insecticide;				
KM	pharmaceutical.				
XX	XX	XX			
OS	Drosophila melanogaster.				
XX	XX	XX			
FN	WO200171042-A2.				
XX	XX	XX			
PD	27-SEP-2001.				
XX	XX	XX			
PF	23-MAR-2001; 2001WO-US009231.				
XX	XX	XX			
PR	23-MAR-2000; 2000US-0191637P.				
PR	11-JUL-2000; 2000US-00614150.				
XX	XX	XX			
PA	(PEKE) PE CORP NY.				
XX	XX	XX			
PI	Venter JC, Adams M, Li PWD, Myers EW;				
XX	XX	XX			
DR	WPI; 2001-656860/75.				
DR	N-PSDB; ABL09265.				
XX	XX	XX			
PT	New isolated nucleic acid detection reagent for detecting 1000 or more				
PT	genes from Drosophila and for elucidating cell signaling and cell-cell				
PT	interactions.				
XX	XX	XX			
PS	Disclosure; SEQ ID NO 22278; 21pp + Sequence Listing; English.				
XX	XX	XX			
CC	The invention relates to an isolated nucleic acid detection reagent				
CC	capable of detecting 1000 or more genes from drosophila. The invention is				
CC	useful in developmental biology and in elucidating cell signalling and				
CC	cell-cell interactions in higher eukaryotes for the development of				
CC	insecticides, therapeutics and pharmaceutical drugs. The invention				
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA				
CC	sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-				
CC	ABB72072). The sequence data for this patent did not form part of the				
CC	printed specification, but was obtained in electronic format directly				
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences				
XX	XX	XX			
SQ	Sequence 567 AA;				
 Query Match Best Local Similarity 100.0%; Pred. No. 7e+02;					
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	15	VLLGGVL 21			
Db	428	VLLGGVL 434			
 RESULT 899					
ADS28277					
ID	ADS28277 standard; protein; 585 AA.				
XX	XX	XX			
AC	ADS28277;				
XX	XX	XX			

RESULT 900
ADH12819
ID ADH12819 standard; protein; 594 AA.
XX AC ADH12819;
XX DT 11-MAR-2004 (first entry)
XX DE Abalone (Haliotis discus hannai) cellulase, SEQ ID NO:1.
XX KW Abalone; cellulase; liver pancreas; de-inking; decolouring;
KW paper manufacture; detergent; food manufacture; animal fodder;
KW biomass degradation; energy production; cellico-oligosaccharide production;
KW cellobiose; cellulose; waste recycling; EC 3.2.1.4; enzyme.
XX OS Haliotis discus; subsp. hannai.
XX PH Key Location/Qualifiers
FT Peptide 1..15
FT Protein /label= Signal_peptide
FT 16..594
FT /note= "Mature cellulase"
PN JP2003235552-A.
XX 26-AUG-2003.
PD
XX 13-FEB-2002; 2002JP-00034852.
PF
XX 13-FEB-2002; 2002JP-00034852.
PR
XX (HOKK-) HOKKAIDO TLO KK.
PA WPI; 2004-147477/15.
DR N-PSDB; ADH12800.
XX
XX Novel cellulase originating in spiral shells capable of degrading
PT cellulose, useful for producing cell oligosaccharide such as cellobiose
PT and a cellulose.
XX
XX Claim 1; SEQ ID NO 1; 21pp; Japanese.
XX
XX The invention relates to a cellulase (ADH12819) from the abalone Haliotis
CC discus hannai. The cellulase is present in high levels in the liver
CC pancreas of the abalone, particularly the anterior portion. The mature
CC cellulase (EC 3.2.1.4) has an N-terminal sequence given in ADH12801, and
CC has a molecular weight of 66 kD, a pH optimum of 5.5-8.0, an optimum
CC temperature of 35-40 degrees Celsius, and is stable at 40 degrees Celsius
CC or below. The invention also encompasses a method for the preparation of
CC the cellulase from abalone internal organs, and further discloses a cDNA
CC sequence (ADH12800) encoding the cellulase. The abalone cellulase is
CC useful in the de-inking and decolouring processes in paper manufacturing,
CC in detergent compositions, and in the manufacture of foods or animal
CC fodder. It is also useful in the degradation of biomass for energy
CC production and for the production of cellico-oligosaccharides such as
CC cellobiose and cellulose. The cellulase of the invention is extracted
CC from parts of abalone which are inedible, thousands of tonnes of which
CC are discarded during processing of the shellfish as a foodstuff. The
CC method of the invention permits helps to reduce the environmental impact
CC of this waste, and permits the inexpensive and rapid production of large
CC quantities of cellulase. The present sequence represents the abalone
CC cellulase of the invention. Note: The present sequence is listed as SEQ
CC ID NO:1 along with the cDNA encoding it in the sequence listing. However,
CC claim 1 refers to the protein sequence only of SEQ ID NO:1.
XX
XX Sequence 594 AA;
SQ
Query Match 5.9%; Score 7; DB 8; Length 594;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
20 VLAALAA 26

Db 5 VLAALAA 11

RESULT 901

ADN19948
ID ADN19948 standard; protein; 605 AA.
XX AC ADN19948;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #2601.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX (CAOY/) CAO Y.
PA (HINKLE/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 2601; 122pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 605 AA;
Query Match 5.9%; Score 7; DB 8; Length 605;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 84 GKVLGILL 90
DB 389 GKVLGILL 395
RESULT 902
ADSD29014
ID ADS29014 standard; protein; 607 AA.
XX AC ADS29014;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #18047.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 607 AA;
Query Match 5.9%; Score 7; DB 8; Length 607;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALAA 26
DB 408 VLAALAA 414
RESULT 903
ADSD21496
ID ADS21496 standard; protein; 607 AA.
XX AC ADS21496;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #10529.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX CC New recombinant DNA construct comprising a promoter positioned to provide
XX CC for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source, useful for producing plants with improved properties.
XX CC Claim 1; SEQ ID NO 10529; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 607 AA;

Query Match 5.9%; Score 7; DB 8; Length 607;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
 |||||
 Db 408 VLALAA 414

RESULT 904
 ABP64821
 ID ABP64821 standard; protein; 622 AA.

AC ABP64821;

DT 25-FEB-2003 (first entry)

XX Human protein SEQ ID 481.

XX Human; expressed sequence tag; EST; haematopoietic disorder;
 KW central nervous system disease; viral infection;
 KW peripheral nervous system disease; non-healing wound; infectious disease;
 KW immune deficiency; immune disorder; bacterial infection; allergy; cancer;
 KW fungal infection; autoimmune disorder; coagulation disorder; neutropenic;
 KW antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
 KW cytosstatic; haemostatic; virucide; antibacterial; fungicide;
 KW immunostimulant; cerebroprotective.

XX Homo sapiens.

XX WO200259260-A2.

XX 01-AUG-2002.

XX 16-NOV-2001; 2001WO-US042950.

XX 17-NOV-2000; 2000US-00714936.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
 PI Ren F, Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-590824/63.

XX N-PSDB; ABQ99407.

XX New isolated polynucleotide, useful in research, diagnostic or
 PT therapeutic methods, e.g. preventing or treating disorders involving
 PT aberrant protein expression or biological activity.

XX Claim 20; SEQ ID NO 481; 394pp; English.

XX The present invention relates to novel human coding sequences (ABQ99268-
 CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
 CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production

CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotential or pluripotential state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 622 AA;

Query Match 5.9%; Score 7; DB 5; Length 622;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 |||||
 Db 155 GGVLAAL 161

RESULT 905

ADQ95978
 ID ADQ95978 standard; protein; 622 AA.

XX ADQ95978;

XX 07-OCT-2004 (first entry)

XX T cell activation associated protein #78.

XX antiallergic; antiarthritic; antiaesthetic; antidiabetic; anti-HIV;
 KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
 KW gene therapy; T cell activation; diagnosis; autoimmune disease;
 KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
 KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
 KW bone-marrow transplant.

XX Homo sapiens.

XX WO2004058805-A2.

XX 15-JUL-2004.

XX 25-DEC-2003; 2003WO-JP016715.

XX 26-DEC-2002; 2002JP-00376365.

XX 27-DEC-2002; 2002US-0436473P.

XX 25-APR-2003; 2003JP-00122113.

XX 28-APR-2003; 2003US-0465792P.

XX 21-OCT-2003; 2003JP-00360559.

XX 22-OCT-2003; 2003US-0512846P.

XX (ASAH-) ASahi Kasei Pharma Corp.

XX Matsuda A, Yoneta S;

XX WPI; 2004-593134/57.

XX N-PSDB; ADQ95977.

XX New purified protein involved in T cell activation, useful for

PT diagnosing, preventing and/or treating acquired immunodeficiency
PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic
PT and infectious diseases.
XX
XX
XX Claim 1; SEQ ID NO 156; 2828pp; English.
XX
XX The invention relates to purified proteins and genes encoding them, that
CC are involved in T cell activation (1) and has an amino acid deletion,
CC substitution or addition in the amino acid sequences. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
CC marrow transplant. This sequence corresponds to a protein involved in T
CC cell activation.
XX
XX
SQ Sequence 622 AA;

Query Match 5.9%; Score 7; DB 8; Length 622;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
Db 155 GGVLAAL 161
|||||
|
RESULT 906
AAU74619
ID AAU74619 standard; protein; 626 AA.
XX
AC AAU74619;
XX
DT 09-APR-2002 (first entry)
XX
DE Oestrogen-regulated LIV-1 family protein BAA91091_Hs.
XX
XX LIV-1; oestrogen; cytostatic; neuroprotective; zinc homeostasis;
XX gene therapy; apoptosis modulator; cancer; neurodegenerative disorder;
XX apoptotic disorder; zinc-homeostasis related disorder.
XX
XX Homo sapiens.
XX
XX WO200196372-A2.
XX
XX 20-DEC-2001.
XX
XX 13-JUN-2001; 2001WO-GB002597.
XX
XX 13-JUN-2000; 2000GB-00014411.
XX
XX 14-JUN-2000; 2000GB-00014493.
XX
XX 05-JUL-2000; 2000US-0216349P.
XX
XX (UYCA-) UNIV COLLEGE CARDIFF.
XX
XX Taylor KM, Morgan HE, Nicholson RI;
XX WPI; 2002-106465/14.
XX
XX Use of a polypeptide comprising one or more consensus regions of proteins
PT of LIV-1 family for treating disorders of zinc homeostasis such as breast
PT cancer, neurodegenerative disorders, and for modifying apoptosis.
XX
XX Claim 1; Fig 1; 67pp; English.
XX
XX The invention describes the a polypeptide comprising one or more
CC consensus regions of proteins of LIV-1 family or its functional
CC homologue. The polypeptide is useful in the preparation of a medicament
CC for the treating a disease e.g. those involving disorders of zinc
CC homeostasis, in gene therapy and for modifying apoptosis in vitro or in
CC vivo on contact with cells. Diseases involving defects in zinc
CC homeostasis include cancer, neurodegenerative disorders and apoptotic
CC disorders. Recombinant proteins of the LIV-1 family (an oestrogen-

CC regulated gene) are useful for diagnosing a zinc homeostasis-related
CC condition in a subject which involves contacting a sample from the
CC subject with the recombinant protein and measuring the binding of
CC antibody to the sample. The antibody is also useful for treating a zinc
CC homeostasis-related condition. This sequence is a member of the LIV-1
CC family (a gene regulated by oestrogen levels) and is useful for creating
CC recombinant proteins for diagnosing zinc-homeostasis related conditions,
XX described in the method of the invention
XX
SQ Sequence 626 AA;

Query Match 5.9%; Score 7; DB 5; Length 626;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
Db 155 GGVLAAL 161
|||||
|
RESULT 907
ABG96322
ID ABG96322 standard; protein; 626 AA.
XX
AC ABG96322;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M439.
XX
XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;
XX central nervous system disorder; bacterial meningitis; viral meningitis;
XX Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
XX brain herniation; inflammation; encephalitis; testicular disorder;
XX non-tuberculous granulomatous orchitis; connective tissue disorder;
XX heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
XX histological type; carcinogenic; ovarian cancer marker.
XX
XX Homo sapiens.
XX
XX WO200271928-A2.
XX
XX 19-SEP-2002.
XX
XX 14-MAR-2002; 2002WO-US007826.
XX
XX 14-MAR-2001; 2001US-0276025P.
XX
XX 14-MAR-2001; 2001US-0276026P.
XX
XX 10-AUG-2001; 2001US-0311732P.
XX
XX 19-SEP-2001; 2001US-0323580P.
XX
XX 26-SEP-2001; 2001US-0324967P.
XX
XX 26-SEP-2001; 2001US-0325102P.
XX
XX 26-SEP-2001; 2001US-0325149P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
XX Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
XX Baat RC, Lu K, Schmandt RE, Zhao X, Glatt K;
XX WPI; 2002-732277/78.
XX
XX N-PSDB; ABS76414.
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
XX Disclosure; Page 222-223; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level

of expression of the marker in a control non-ovarian cancer sample, where the marker is selected from 363 cancer markers described in the specification. The method of the invention is useful in diagnosing or characterising cancer, in detecting the presence of cancer as early as possible, and the recurrence of ovarian cancer. The method may also be of particular use with patients having an enhanced risk of developing ovarian cancer (e.g. patients having a familial history of ovarian cancer). The cancer markers may be used in the management and treatment of e.g. brain and central nervous system disorders (e.g. bacterial and viral meningitis, Alzheimer's disease or Parkinson's disease), brain disorders (e.g. cerebral oedema, hydrocephalus or brain herniations), inflammations (e.g. bacterial or viral meningitis or encephalitis), testicular disorders (e.g. nontuberculous granulomatous orchitis), connective tissue disorders, or heart disorders (e.g. ischaemic heart disease or atherosclerosis). The compositions and methods may also be used in assessing the histological type of neoplasm associated with ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian carcinogenic potential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer. The present amino acid sequence represents one of the ovarian cancer markers described in the invention

XX SQ Sequence 626 AA;

Query Match 5.9%; Score 7; DB 5; Length 626;

Best Local Similarity 100.0%; Pred. No. 7.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24

Db 155 GGVLAAL 161

RESULT 908

ABP70108

ID ABP70108 standard; protein; 626 AA.

XX AC ABP70108;

XX DT 27-JAN-2003 (first entry)

XX DE Human NOV26a.

XX KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS; antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic; nontropic; immunosuppressive; osteopathic; antiparkinsonian; cancer; antiinfertility; cerebroprotective; gene therapy; NOVA; NOV; fertility; metabolic disorder; diabetes; obesity; infectious disease; anorexia; neurodegenerative disease; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorder; cardiovascular disorder; bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis; KW metabolic syndrome X; wasting disorder; cell differentiation; cell proliferation; haematopoiesis; wound healing; angiogenesis.

XX OS Homo sapiens.

XX PN WO200272771-A2.

XX PD 19-SEP-2002.

XX PF 08-MAR-2002; 2002WO-US007288.

XX PR 08-MAR-2001; 2001US-0274101P.

XX PR 08-MAR-2001; 2001US-0274194P.

XX PR 08-MAR-2001; 2001US-0274281P.

XX PR 09-MAR-2001; 2001US-0274322P.

XX PR 12-MAR-2001; 2001US-0274849P.

XX PR 13-MAR-2001; 2001US-0275235P.

XX PR 13-MAR-2001; 2001US-0275578P.

XX PR 13-MAR-2001; 2001US-0275579P.

XX PR 13-MAR-2001; 2001US-0275601P.

PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277321P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 28-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 04-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291099P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 12-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 03-DEC-2001; 2001US-0337426P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.

(CURA-) CURAGEN CORP.

Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CM;
Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VV, Pochart PF;
Zhong M;

WPI; 2002-732824/79.

N-ESDB; ABV99386.

New NOVX polypeptides and polynucleotides, useful for preventing, diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer, Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic disorders, and asthma.

Claim 1; Page 188; 619pp; English.

```

XX
CC The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome
CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, haematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX
SQ Sequence 626 AA;

Query Match 5.9%; Score 7; DB 5; Length 626;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
DB 155 GGVLAAL 161
|||||

RESULT 909
ADU06701
ID ADU06701 standard; protein; 626 AA.
XX AC ADU06701;
XX DT 27-JAN-2005 (first entry)
XX DE Novel bronchial cancer-associated human protein SeqID927.
XX KW bronchial cancer; cytostatic; tumour-associated protein;
XX cancer detection; metastasis; tumour; human.
XX OS Homo sapiens.
XX PN DE10316701-Al.
XX PD 04-NOV-2004.
XX PF 09-APR-2003; 2003DE-01016701.
XX PR 09-APR-2003; 2003DE-01016701.
XX PA (HINZ/) HINZMANN B.
XX PA (HERM/) HERMANN K.
XX PA (CAST/) HEIDEN CASTANOS-VELEZ E.
XX
XX Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;
XX Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarsky C;
XX
XX WPI; 2004-786403/78.
XX DR N-PSDB; ADU06214.
XX
XX New nucleic acid, and derived proteins, useful for diagnosis of bronchial
XX cancer and in screening for therapeutic and diagnostic agents.
XX
XX Claim 2; SEQ ID NO 927; 1381pp; German.
XX
XX This invention relates to a novel isolated nucleic acid associated with
XX bronchial cancer comprising 489 defined sequences given in the
XX specification. The invention may be useful for the production of
XX

```

CC	compounds with a cytostatic activity through the inhibition of expression
CC	or activity of tumour-associated proteins. The novel DNA sequences and
CC	the proteins/peptides encoded by them are used for detecting bronchial
CC	cancer or determining the risk of developing it and to screen for
CC	specific binding partners of the DNA or protein sequences, where the
CC	binding partners are potentially useful as agents for treating or
CC	diagnosing bronchial cancer. The DNA or protein sequences can also be
CC	used for prognosis, detection of metastases and for secondary treatment
CC	(of tumours that have been stabilised or are no longer detectable).
CC	Detecting abnormal expression of the DNA sequences provides early
CC	diagnosis of bronchial cancers. The present sequence is that of a protein
CC	encoded by a novel bronchial cancer-associated human gene sequence of the
CC	invention.
XX	
SQ	Sequence 626 AA;
	Query Match 5.9%; Score 7; DB 8; Length 626;
	Best Local Similarity 100.0%; Pred. No. 7.6e+02;
	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	18 GGVLAAL 24
Db	155 GGVLAAL 161
RESULT 910	
AAE06574	
ID	AAE06574 standard; protein; 647 AA.
XX	
AC	AAE06574;
XX	
DT	25-SEP-2001 (first entry)
XX	
DE	Human protein having hydrophobic domain, HP10755.
XX	
KW	Human; hydrophobic domain; gene therapy; nutritional supplement;
KW	cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
KW	multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
KW	haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
KW	Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
KW	haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
KW	contraceptive; antiinfertility; antinflammatory.
XX	
OS	Homo sapiens.
XX	
PN	W0200149728-A2.
XX	
PD	12-JUL-2001.
XX	
PF	28-DEC-2000; 2000WO-JP009359.
XX	
PR	06-JAN-2000; 2000JP-00000585.
PR	06-JAN-2000; 2000JP-00000588.
PR	11-JAN-2000; 2000JP-00002299.
PR	03-FEB-2000; 2000JP-00026862.
PR	03-MAR-2000; 2000JP-00058367.
XX	
PA	(PROT-) PROTEGENE INC.
PA	(SAGA) SAGAMI CHEM RES CENT.
XX	
PI	Kato S, Kimura T;
XX	
DR	WPI; 2001-418355/44.
DR	N-PSDB; AAD12569.
XX	
PT	Human proteins with hydrophobic domains and the nucleic acids encoding
PT	them, useful for preventing diagnosing and treating e.g. cancer,
XX	Alzheimer's and inflammation.
XX	
PS	Claim 1; Page 67-68; 563pp; English.
XX	
CC	The present sequence is human protein with hydrophobic domain, HP10755.
CC	The polynucleotide and polypeptide of the invention may be used in the

CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate polypeptide expression. The polynucleotides may be used to
 CC produce the polypeptide, by inserting the nucleic acids into a host cell
 CC and culturing the cell to express the protein. The polynucleotides and
 CC its complementary sequences may also be used as DNA probes in diagnostic
 CC assays and also used in gene therapy. The polypeptides may also be used
 CC as antigens in the production of antibodies and in assays to identify
 CC modulators of polypeptide expression and activity. The polypeptides and
 CC nucleic acids may be used as nutritional supplements, to modulate
 CC cytokine and cell proliferation activity, to modulate immune stimulation
 CC or suppression (e.g. for the treatment of microbial infections and
 CC autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and
 CC insulin-dependent diabetes), to modulate haematopoiesis, to modulate
 CC tissue growth activity (e.g. for the treatment of Parkinson's disease,
 CC Huntington's disease and Alzheimer's disease), to modulate activin and
 CC inhibin activity (e.g. for controlling fertility), to modulate
 CC chemotactic and chemokinetic activity, to modulate haemostatic and
 CC thrombolytic activity, to modulate receptor ligand activity, to modulate
 CC inflammation and to inhibit tumour growth
 XX
 SQ Sequence 647 AA;

Query Match 5.9%; Score 7; DB 4; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 |||||
 Db 180 GGVLAAL 186

RESULT 911
 AAB74710
 ID AAB74710 standard; protein; 647 AA.

AC AAB74710;

DT 12-JUN-2001 (first entry)

DE Human membrane associated protein MEMAP-16.

XX Human; membrane associated protein; MEMAP; diagnosis; cytostatic;
 KW antiinflammatory; anticonvulsant; immunosuppressive; antidiarrheic;
 KW antiarteriosclerotic; gene therapy; cell proliferative disorder;
 KW autoimmune disorder; inflammatory disorder; neurological disorder;
 KW gastrointestinal disorder; cancer; inflammation; atherosclerosis;
 KW epilepsy; diarrhoea.

XX Homo sapiens.

FN WO200112662-A2.

PD 22-FEB-2001.

PF 14-AUG-2000; 2000WO-US022315.

PR 17-AUG-1999; 99US-0149641P.

PR 09-NOV-1999; 99US-0164203P.

XX (INCY-) INCYTE GENOMICS INC.

PI Lal P, Yue H, Tang YT, Bandman O, Burford N, Azimzai Y;

PI Baughn MR, Lu DAM, Patterson C;

XX WPI; 2001-168860/17.

DR N-PSDB; AAF81756.

XX Isolated polypeptide with a human membrane associated protein sequence is
 XX useful for the diagnosis, prevention and treatment of cell proliferative,
 XX autoimmune/inflammatory, neurological and gastrointestinal disorders.

PS Claim 1; Page 128-130; 173pp; English.

XX

CC AAF81741 to AAF81777 encode the human membrane associated proteins
 CC (MEMAP) given in AAB74695 to AAB74731. MEMAPs have cytostatic,
 CC antiinflammatory, anticonvulsant, immunosuppressive, antidiarrheic and
 CC antiarteriosclerotic activities, which can be used in gene therapy.
 CC MEMAPs and agonist of MEMAPs can be used to treat a disease or condition
 CC associated with decreased expression of functional MEMAP and antagonists
 CC of MEMAP are used to treat a disease or condition associated with
 CC overexpression of functional MEMAP. These disorders include cell
 CC proliferative, autoimmune/inflammatory, neurological and gastrointestinal
 CC disorders. The MEMAP polynucleotides and proteins are also used for the
 CC diagnosis of these disorders. Specific examples of these disorders
 CC include cancer, inflammation, atherosclerosis, epilepsy and diarrhoea.
 CC MEMAP proteins can be used to screen for compounds which specifically
 CC bind MEMAP including antibodies, oligonucleotides, proteins and small
 CC molecules. MEMAP polynucleotides can be used to prepare transgenic
 CC animals which can be studied to provide information concerning human
 CC disease. Anti-MEMAP antibodies are useful in immunoassays for the
 CC detection of MEMAP protein and can be used as antagonists to treat or
 CC prevent a disorder associated with MEMAP. Polynucleotides encoding MEMAP
 CC can be delivered to target cells with genetic abnormalities with respect
 CC to the expression of MEMAP to treat or prevent a disorder associated with
 CC MEMAP

XX Sequence 647 AA;

Query Match 5.9%; Score 7; DB 4; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 |||||
 Db 180 GGVLAAL 186

RESULT 912

AAB01677

ID AAB01677 standard; protein; 647 AA.

AC AAB01677;

DT 18-JUL-2001 (first entry)

XX Human gene 6 encoded secreted protein HWHM66, SEQ ID NO:89.

XX Human; secreted protein; proliferative disorder; cancer; tumour; asthma;
 KW fetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
 KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
 KW inflammation; neurological disorder; Alzheimer's disease; food additive;
 KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
 KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
 KW cell culture; chemotaxis; vulnerability; binding partner identification;
 KW gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22

FT /label= Signal_peptide

FT Protein 23..647

FT /label= Mature_human_secreted_protein

XX WO200134767-A2.

XX 17-MAY-2001.

XX 01-NOV-2000; 2000WO-US030036.

XX 05-NOV-1999; 99US-0163576P.

PR 27-JUL-2000; 2000US-0221366P.

XX (HUMA-) HUMAN GENOME SCI INC.

```
XX Soppet DR, Komatsoulis G, Shi Y, Oleen HS, Ruben SM;
PI WPI; 2001-316492/33.
DR N-PSDB; AAD05497.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX
XX Claim 11; Page 467-469; 540pp; English.
XX
XX RAD05492-AAD05564 represent cDNAs corresponding to 22 human secreted
CC protein genes, and AAE01672-AAE01743 represent the proteins they encode.
CC AAE01744-AAE01763 represent human secreted protein fragments or variants.
CC The secreted proteins and their genes are useful for preventing, treating
CC or ameliorating medical conditions, e.g., by protein or gene therapy.
CC Pathological conditions can be diagnosed by determining the amount of the
CC new protein in a sample or by determining the presence of mutations in
CC the new genes. Specific uses are described for each of the 22 genes,
CC based on the tissues in which they are most highly expressed, and include
CC developing products for the diagnosis or treatment of proliferative
CC disorders, cancer, tumours, foetal and developmental abnormalities,
CC haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
CC diseases (e.g., rheumatoid arthritis), inflammation, allergies,
CC neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
CC cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
CC psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
CC angiogenic disorders, kidney disorders, gastrointestinal disorders,
CC pregnancy-related disorders, endocrine disorders, and infections. The
CC proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues,
CC to regenerate tissues, to identify their cognate ligands or binding
CC partners, and in chemotaxis, and can be used as a food additive or
CC preservative to modify storage properties. Antibodies specific for a
CC protein of the invention can be used in alleviating symptoms associated
CC with the disorders mentioned above, and in diagnostic immunoassays e.g.,
CC radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The
CC present sequence represents a human secreted protein of the invention
XX
XX Sequence 647 AA;
XX
XX Query Match 5.9%; Score 7; DB 4; Length 647;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAAL 24
XX | | | | |
XX Db 180 GGVLAAL 186
XX
XX RESULT 913
XX AAU74618
XX ID AAU74618 standard; protein; 647 AA.
XX
XX AC AAU74618;
XX
XX DT 09-APR-2002 (first entry)
XX
XX DE Oestrogen-regulated LIV-1 family protein AX083511_Hs.
XX
XX KW LIV-1; oestrogen; cytostatic; neuroprotective; zinc homeostasis;
XX gene therapy; apoptosis modulator; cancer; neurodegenerative disorder;
XX apoptotic disorder; zinc-homeostasis related disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200196372-A2.
XX
XX PD 20-DEC-2001.
XX
XX PF 13-JUN-2001; 2001WO-GB002597.
XX
XX
```

```
PR 13-JUN-2000; 2000GB-00014411.
PR 14-JUN-2000; 2000GB-00014493.
PR 05-JUL-2000; 2000US-0216349P.
XX
XX (UYCA-) UNIV COLLEGE CARDIFF.
XX
XX Taylor KM, Morgan HE, Nicholson RI;
XX
XX WPI; 2002-106465/14.
XX
XX Use of a polypeptide comprising one or more consensus regions of proteins
PT of LIV-1 family for treating disorders of zinc homeostasis such as breast
PT cancer, neurodegenerative disorders, and for modifying apoptosis.
XX
XX Claim 1; Fig 1; 67pp; English.
XX
XX The invention describes the a polypeptide comprising one or more
CC consensus regions of proteins of LIV-1 family or its functional
CC homologue. The polypeptide is useful in the preparation of a medicament
CC for the treating a disease e.g. those involving disorders of zinc
CC homeostasis, in gene therapy and for modifying apoptosis in vitro or in
CC vivo on contact with cells. Diseases involving defects in zinc
CC homeostasis include cancer, neurodegenerative disorders and apoptotic
CC disorders. Recombinant proteins of the LIV-1 family (an oestrogen-
CC regulated gene) are useful for diagnosing a zinc homeostasis-related
CC condition in a subject which involves contacting a sample from the
CC subject with the recombinant protein and measuring the binding of a
CC antibody to the sample. The antibody is also useful for treating a zinc
CC homeostasis-related condition. This sequence is a member of the LIV-1
CC family (a gene regulated by oestrogen levels) and is useful for creating
CC recombinant proteins for diagnosing zinc-homeostasis related conditions,
CC described in the method of the invention
XX
XX Sequence 647 AA;
XX
XX Query Match 5.9%; Score 7; DB 5; Length 647;
XX Best Local Similarity 100.0%; Pred. No. 7.8e+02;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 18 GGVLAAL 24
XX | | | | |
XX Db 180 GGVLAAL 186
XX
XX RESULT 914
XX ABG63939
XX ID ABG63939 standard; protein; 647 AA.
XX
XX AC ABG63939;
XX
XX DT 27-AUG-2002 (first entry)
XX
XX DE Human albumin fusion protein #614.
XX
XX KW Albumin fusion protein; therapeutic protein X; human albumin; HA;
XX human serum albumin; HSA; cancer; reproductive disorder;
XX digestive disorder; immune disorder; endocrine disorder;
XX haematopoietic disorder; neural disorder; connective disorder;
XX cytostatic; antiinfectivity; antinflammatory; antiulcer;
XX immunomodulator; anti-HIV; antidiabetic; haemostatic; neuroleptic;
XX neuroprotective; antiparkinsonian; antimicrobial; neuroleptic;
XX osteopathic; antiarthritic.
XX
XX OS Homo sapiens.
XX
XX PN WO200177137-A1.
XX
XX PD 18-OCT-2001.
XX
XX PF 12-APR-2001; 2001WO-US011988.
XX
XX PF 12-APR-2000; 2000US-0229358P.
XX
```

PR 25-APR-2000; 2000US-0199384P.
 PR 21-DEC-2000; 2000US-0256931P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Rosen CA, Haeeltine WA;
 XX WPI; 2002-010886/01.
 DR New fusion protein for treating disease e.g. diabetes comprises an
 XX albumin fused to a therapeutic protein.
 XX Claim 1; Page 916-918; 2102pp; English.
 XX The present invention relates to albumin fusion proteins comprising a
 CC therapeutic protein X and human albumin (HA, also known as human serum
 CC albumin, HSA). The proteins are useful for treating a disease or disorder
 CC that may be modulated by therapeutic protein X. The albumin extends the
 CC shelf-life of protein X, and may increase its biological in vitro/in vivo
 CC activity. The protein is useful for treating and diagnosing disorders
 CC such as cancer, reproductive disorders, digestive disorders (e.g. Crohn's
 CC disease, ulcerative colitis), immune disorders (e.g. acquired
 CC immunodeficiency syndrome, AIDS), endocrine disorders (e.g. diabetes),
 CC haematopoietic disorders, neural disorders (e.g. Alzheimer's,
 CC Parkinson's, Creutzfeldt-Jacob disease, encephalomyelitis, meningitis,
 CC schizophrenia), and connective disorders (e.g. osteoporosis, arthritis).
 CC ABG63326-ABG65518 represent albumin fusion proteins of the invention
 XX
 SQ Sequence 647 AA;
 Query Match 5.9%; Score 7; DB 5; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAAL 24
 Db 180 GGVLAAL 186
 RESULT 915
 ADE31719
 ID ADE31719 standard; protein; 647 AA.
 AC ADE31719;
 XX 29-JAN-2004 (first entry)
 DE Human 64624 protein #SEQ ID 76.
 XX Antiarteriosclerotic; cardiac; vasotropic; antiinflammatory;
 KW thrombolytic; antiarrhythmic; antianginal; hypotensive; gene therapy;
 KW cardiovascular; disorder; ischaemia; aortic bending;
 KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;
 KW angina; cardiomyopathy; cardiac death.
 XX Homo sapiens.
 OS
 XX WO2003065984-A2.
 PN
 XX 14-AUG-2003.
 PD
 XX 29-JAN-2003; 2003WO-US002571.
 XX 01-FEB-2002; 2002US-0353224P.
 PR 15-MAR-2002; 2002US-0364529P.
 PR 19-APR-2002; 2002US-0373861P.
 PR 29-APR-2002; 2002US-0376287P.
 PR 12-JUN-2002; 2002US-0388080P.
 PR 24-JUN-2002; 2002US-0390971P.
 PR 03-JUL-2002; 2002US-0394130P.
 PR 10-JUL-2002; 2002US-0394797P.
 PR 21-AUG-2002; 2002US-0404904P.
 PR 23-AUG-2002; 2002US-0405450P.

PR 04-SEP-2002; 2002US-0408070P.
 PR 06-NOV-2002; 2002US-0424300P.
 PR 05-DEC-2002; 2002US-0431042P.
 PR 05-DEC-2002; 2002US-0431079P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Logan TJ, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
 XX Stagliano N, Perodin J, Rodrigue-Way A;
 XX WPI; 2003-731468/69.
 DR N-PSDB; ADE31718.
 XX
 PT Identifying a compound capable of treating a cardiovascular disorder
 PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
 PT modulate the expression or activity of e.g. 1682, 6169 or 6193
 XX polypeptide or nucleic acid.
 PS Disclosure; SEQ ID NO 76; 328pp; English.
 XX The invention relates to a method for identifying a compound capable of
 CC treating a cardiovascular disorder. The present invention identifies the
 CC differential expression of 1682, 6169, 6193, 7771, 14395, 29002, 33216,
 CC 43726, 69292, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,
 CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,
 CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
 CC 2868, 283, 2354, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
 CC 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or
 CC 6585 genes in cardiovascular disease states. The methods are useful in
 CC diagnosing, preventing and treating cardiovascular disorders, such as
 CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,
 CC restenosis, arterial inflammation, vascular wall remodeling, coronary
 CC microembolism, tachycardia, bradycardia, pressure overload, aortic
 CC bending, coronary artery ligation, vascular heart disease, valvular
 CC disease, including but not limited to, valvular degeneration caused by
 CC calcification, rheumatic heart disease, endocarditis, or complications of
 CC artificial valves; atrial fibrillation, long-QT syndrome, congestive
 CC heart failure, sinus node dysfunction, angina, heart failure,
 CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,
 CC including but not limited to, pericardial effusion and pericarditis;
 CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
 CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary
 CC artery spasm, ischaemic disease, arrhythmia, sudden cardiac death and
 CC cardiovascular developmental disorders. The methods may also be used for
 CC identifying compounds that modulate cardiovascular disorders. Sequences
 CC given in ADE31644-ADE31769 represent the genes and proteins that may be
 CC regulated by a compound of the invention.
 XX
 SQ Sequence 647 AA;
 Query Match 5.9%; Score 7; DB 7; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAAL 24
 Db 180 GGVLAAL 186
 RESULT 916
 ADN40019
 ID ADN40019 standard; protein; 647 AA.
 XX
 XX ADN40019;
 XX 17-JUN-2004 (first entry)
 DT
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C389.
 XX Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;

KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
 KW vulnery; gene therapy; vaccine.

XX Homo sapiens.

XX WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

XX 21-NOV-2001; 2001US-0332464P.

XX 29-NOV-2001; 2001US-0334333P.

XX 03-DEC-2001; 2001US-0335339P.

XX 14-DEC-2001; 2001US-0340376P.

XX 08-JAN-2002; 2002US-0347211P.

XX 10-JAN-2002; 2002US-0347349P.

XX 08-FEB-2002; 2002US-0355250P.

XX 13-FEB-2002; 2002US-0356714P.

XX 20-FEB-2002; 2002US-0359077P.

XX 29-MAR-2002; 2002US-0368809P.

XX 04-APR-2002; 2002US-0370110P.

XX 12-APR-2002; 2002US-0372246P.

XX 05-JUN-2002; 2002US-0386614P.

XX 16-JUN-2002; 2002US-0396839P.

XX 22-JUL-2002; 2002US-0397775P.

XX 20-JUL-2002; 2002US-0397845P.

XX 09-SEP-2002; 2002US-0409450P.

XX (E0SB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI; 2003-468649/44.

XX N-PSDB; ADN39802.

XX Determining the presence or absence of a pathological cell in a patient,
 XX useful for diagnosing, prognosing or treating cancer, comprises detecting
 XX a nucleic acid in a biological sample.

XX Claim 12; SEQ ID NO C389; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 XX whose expression is upregulated or downregulated in specific cancers or
 XX other diseases such as angiogenic or fibrotic disorders, and to methods
 XX of determining the presence or absence of a pathological cell in a
 XX patient by detecting a nucleic acid at least 80% identical to those of
 XX the invention or by detecting a polypeptide of the invention. The
 XX invention also relates to expression vectors and host cells comprising a
 XX nucleic acid of the invention; antibodies which specifically bind a
 XX polypeptide of the invention; use of such antibodies for drug targeting;
 XX and methods of screening for modulators of activity or expression of the
 XX polypeptides and nucleic acids. The nucleic acids, polypeptides,
 XX antibodies and methods are useful for diagnosing, prognosing and treating
 XX cancer and other conditions such as psoriasis, ischaemia, heart disease,
 XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 XX neovascularization syndromes, scarring and uterine fibroids. They may
 XX also be useful in wound healing and in contraception. The present
 XX sequence represents a polypeptide of the invention.

XX Query Match 5.9%; Score 7; DB 7; Length 647;

XX Best Local Similarity 100.0%; Pred. No. 7.8e+02;

XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 Db 180 GGVLAAL 186

RESULT 917

ADL77204

XX ADL77204 standard; protein; 647 AA.

XX AC ADL77204;

XX 20-MAY-2004 (first entry)

XX Albumin fusion protein related therapeutic protein X, SEQ ID NO 686.

XX albumin fusion protein; cytostatic; antineoplastic; antiarthritic;
 XX antiaesthetic; anti-HIV; immunosuppressive; antiinflammatory;
 XX antipsoriatic; antibacterial; osteopathic; dermatological; antigout;
 XX immunomodulator; antiarrhythmic; cardiant; antiparkinsonian; tranquilizer;
 XX nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
 XX antidiabetic; anabolic; hypertensive; vulnery; gene therapy; cancer;
 XX reproductive system disorder; therapeutic protein.

XX Unidentified.

XX US2004010134-A1.

XX 15-JAN-2004.

XX 12-APR-2001; 2001US-00833245.

XX 12-APR-2000; 2000US-0229358P.

XX 25-APR-2000; 2000US-0199384P.

XX 21-DEC-2000; 2000US-0256931P.

XX (ROSE/) ROSEN C A.

XX (HASE/) HASELTINE W A.

XX Rosen CA, Haseltine WA;

XX WPI; 2004-090519/09.

XX New albumin fusion proteins, useful for diagnosing, treating, preventing
 XX or ameliorating diseases or disorders e.g. cancer, anemia, arthritis,
 XX asthma, inflammatory bowel disease or Alzheimer's disease.

XX Disclosure; SEQ ID NO 686; 279pp; English.

XX The invention relates to a novel albumin fusion protein. The invention
 XX further relates to: a composition comprising the albumin fusion protein
 XX and a pharmaceutical carrier; a kit comprising the composition of the
 XX albumin fusion protein formula; a method of treating a disease or
 XX disorder in a patient comprising the step of administering the albumin
 XX fusion protein; a method of treating a patient with a disease or disorder
 XX that is modulated by therapeutic protein: X, or its fragment or variant;
 XX a method of extending the shelf life of therapeutic protein: X, or its
 XX fragment or variant; a nucleic acid molecule comprising a polynucleotide
 XX sequence encoding the albumin fusion protein; a vector comprising the
 XX nucleic acid molecule of the albumin fusion protein; and a host cell
 XX comprising the nucleic acid molecule of the albumin fusion protein. The
 XX albumin fusion protein and its compositions have the following
 XX activities: cytostatic, antineoplastic, antiarthritic, antidiabetic, anti-
 XX HIV, immunosuppressive, antiinflammatory, antipsoriatic, antibacterial,
 XX osteopathic, dermatological, antigout, immunomodulator, antiarrhythmic,
 XX cardiant, nephrotropic, antiparkinsonian, tranquilizer, uropathic,
 XX neuroprotective, antiparkinsonian, tranquilizer, antidiabetic, anabolic,
 XX hypertensive, and vulnery. The albumin fusion protein nucleic acid may
 XX be used in gene therapy to treat disorders. The albumin fusion protein is
 XX useful for diagnosing, treating, preventing or ameliorating diseases or
 XX disorders comprising indication: Y. The diseases or disorders include:
 XX cancer (e.g. leukemia, colon, bone, breast, liver or lung cancer),
 XX immune or haematopoietic diseases (e.g. anemia, Hodgkin's disease, acute
 XX lymphocytic anaemia, multiple myeloma, arthritis, asthma, AIDS,
 XX autoimmune disease, inflammatory bowel disease, psoriasis or Lyme
 XX disease), reproductive system disorders (e.g. prostatitis, inguinal
 XX hernia, varicocele, penile carcinoma, ovarian adenocarcinoma or Sertoli-
 XX Leydig tumours), musculoskeletal diseases (e.g. giant cell tumours,

CC Paget's disease, systemic lupus erythematosus, gout, muscular dystrophy
 CC or cachexia), cardiovascular disease (e.g. rhabdomyomas, heart disease,
 CC arrhythmia, cardiac arrest, heat valve disease, hypernatraemia or
 CC hyponatraemia), mixed foetal diseases (e.g. foetal alcohol syndrome,
 CC Down's syndrome, Patau syndrome, Turner's syndrome, Apert syndrome or Tay
 CC -Sachs disease), excretory diseases (e.g. urinary incontinence, urinary
 CC tract infections or renal disorders), neural or sensory disease (e.g.
 CC Alzheimer's disease, Parkinson's disease, cerebral malaria, meningitis,
 CC cerebellar ataxia, attention deficit disorder, autism or obsessive
 CC compulsive disorder), respiratory disease (e.g. emphysema, lung cancer or
 CC occupational lung disease), endocrine diseases (e.g. diabetes, Addison's
 CC disease or glomerulonephritis), digestive diseases (e.g. portal
 CC hypertension, irritable bowel disease, gastric atrophy or pancreatitis)
 CC or connective tissue or epithelial diseases (e.g. Crohn's disease,
 CC scleroderma, wound healing or epidermolysis bullosa). This sequence
 CC represents a therapeutic protein X relating to the albumin fusion protein
 CC of the invention. The sequence listing data for this specification was
 CC downloaded from the USPTO website.

XX SQ Sequence 647 AA;

Query Match 5.9%; Score 7; DB 8; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 Db 180 GGVLAAL 186
 |||||

RESULT 918
 ADQ95976
 ID ADQ95976 standard; protein; 647 AA.

XX AC ADQ95976;

XX DT 07-OCT-2004 (first entry)

XX DE T cell activation associated protein #77.

XX KW antiallergic; antiarthritic; antiasthmatic; antidiabetic; anti-HIV;
 KW antimicrobial; antirheumatic; immunosuppressive; neuroprotective;
 KW gene therapy; T cell activation; diagnosis; autoimmune disease;
 KW rheumatoid arthritis; asthma; multiple sclerosis; diabetes;
 KW allergic disease; infectious disease; AIDS; chronic rejection; organ;
 KW bone-marrow transplant.

XX OS Homo sapiens.

XX FN WO2004058805-A2.

XX PD 15-JUL-2004.

XX PF 25-DEC-2003; 2003WO-JP016715.

XX PR 26-DEC-2002; 2002JP-00376365.

XX PR 27-DEC-2002; 2002US-0436473P.

XX PR 25-APR-2003; 2003JP-00122113.

XX PR 28-APR-2003; 2003US-0465792P.

XX PR 21-OCT-2003; 2003JP-00360559.

XX PR 22-OCT-2003; 2003US-0512846P.

XX PA (ASAH-) ASAH KASEI PHARMA CORP.

XX PI Mateuda A, Yoneta S;

XX DR WPI; 2004-593134/57.

XX DR N-PSDB; ADQ95975.

XX PT New purified protein involved in T cell activation, useful for

PT diagnosing, preventing and/or treating acquired immunodeficiency

PT syndrome, autoimmune (e.g. rheumatoid arthritis, and diabetes), allergic

PT and infectious diseases.

XX Claim 1; SEQ ID NO 154; 2828pp; English.

XX The invention relates to purified proteins and genes encoding them, that
 CC are involved in T cell activation (I) and has an amino acid deletion,
 CC substitution or addition in the amino acid sequences. The methods and
 CC compositions of the present invention are useful for the diagnosis,
 CC prevention and/or treatment of autoimmune disease (rheumatoid arthritis,
 CC asthma, multiple sclerosis and diabetes), allergic disease, infectious
 CC disease, AIDS, and acute or chronic rejection at organ transplant or bone
 CC marrow transplant. This sequence corresponds to a protein involved in T
 CC cell activation.

XX SQ Sequence 647 AA;

Query Match 5.9%; Score 7; DB 8; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 Db 180 GGVLAAL 186
 |||||

RESULT 919

ADR46689

ID ADR46689 standard; protein; 647 AA.

XX AC ADR46689;

XX DT 18-NOV-2004 (first entry)

XX DE Cancer-associated protein, SEQ ID 102.

XX KW Cytostatic; Gene Therapy; cancer; human.

XX OS Homo sapiens.

XX PN WO2004073657-A2.

XX PD 02-SEP-2004.

XX PF 19-FEB-2004; 2004WO-US005455.

XX PR 19-FEB-2003; 2003US-0448784P.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Aziz N, Gish KC, Wilson KE, Zlotnik A;

XX DR WPI; 2004-652787/63.

XX DR N-PSDB; ADR46631.

XX Detecting a pathological cell in a patient for diagnosing or treating
 PT cancer by detecting in a biological sample from the patient genes whose
 PT expression are up-regulated or down-regulated in specific cancers.

XX Claim 1; SEQ ID NO 102; 375pp; English.

XX The present invention relates to a method for detecting cancer in a
 CC patient. The method comprises detecting in a biological sample from the
 CC patient a nucleotide or protein sequence comprising a sequence that is at
 CC least 80% identical to a nucleotide sequence (ADR46588-ADR46645) or
 CC protein sequence (ADR46646-ADR46703). The method is useful for detecting
 CC cancer for preparing a composition for diagnosing or treating cancer.

XX SQ Sequence 647 AA;

Query Match 5.9%; Score 7; DB 8; Length 647;
 Best Local Similarity 100.0%; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24

```

Db      180 GGVLAAL 186
|||||
RESULT 920
ADZ51373
ID ADZ51373 standard; protein; 647 AA.
XX
AC ADZ51373;
XX
XX 30-JUN-2005 (first entry)
XX
XX Amino acid sequence of ovarian cancer marker M714.
XX
XX cytostatic; gene therapy; ovarian cancer; ovarian cancer marker; M714;
KW KW solute carrier family 39 protein.
XX
XX Homo sapiens.
XX
XX WO2005034732-A2.
PN
XX
XX 21-APR-2005.
PD
XX
XX 07-OCT-2004; 2004WO-US033166.
PF
XX
XX 07-OCT-2003; 2003US-0509171P.
PR
XX
XX (MILL-) MILLENNIUM PHARM INC.
PA
XX
XX Endege WO, Ford D, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S;
PI Monahan JE, Schlegel R, Xu YY, Zhao X;
PI
XX WPI; 2005-306219/31.
DR
DR N-PSDB; ADZ51372.
XX
XX Assessing whether a patient is afflicted with ovarian cancer comprises
PT determining a significant difference between the levels of expression of
PT an ovarian cancer marker in the patient sample and the sample from a
PT control subject.
XX
XX Example 1; SEQ ID NO 30; 164pp; English.
XX
XX The specification describes a method of assessing whether a patient is
CC afflicted with ovarian cancer. The method comprises determining the
CC presence of a significant difference between the levels of expression of
CC an ovarian cancer marker in the patient sample and the sample from a
CC control subject. The method of the invention is useful for assessing,
CC diagnosing, preventing or treating ovarian cancer. The markers may also
CC be used in screening for agents that may treat or prevent ovarian cancer.
CC The present sequence represents the marker M714, which is solute carrier
CC family 39 protein.
XX
XX
XX Sequence 647 AA;
SQ
Query Match 5.9%; Score 7; DB 9; Length 647;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 180 GGVLAAL 186
|||||
RESULT 921
ABZ42807
ID ABZ42807 standard; protein; 647 AA.
XX
AC ABZ42807;
XX
XX 22-SEP-2005 (first entry)
DT
XX
XX Snail transcription factor-related protein SEQ ID NO:34.
XX
KW inflammation; wound healing; vulnery; injury; cancer; cytostatic;
KW neoplasm; epithelial mesenchymal transition inducer; EMT inducer; snail;
KW antiinflammatory; transcription factor.
XX
XX Unidentified.
OS
XX WO2005063301-A1.
PN
XX
XX 14-JUL-2005.
PD
XX
XX 22-DEC-2004; 2004WO-JP019246.
PF
XX
XX 26-DEC-2003; 2003JP-00435122.
PR
XX
XX (HIRA/) HIRANO T.
PA
XX
XX Hirano T, Yamashita S;
PI
XX WPI; 2005-497837/50.
DR
DR N-PSDB; AEB42806.
DR
XX
XX Novel snail active regulator or epithelial mesenchymal transition inducer
PT comprising LIV1 DNA, useful for preparation of wound healing agent.
PT
XX
XX Claim 1; SEQ ID NO 34; 52pp; Japanese.
PS
XX
XX The invention relates to a snail (zinc finger protein) active regulator
CC (I) or epithelial mesenchymal transition (EMT) inducer (II) and encoding
CC polynucleotides (III). Also described are: a snail activation inhibitor
CC (AI) or EMT inhibitor (EI), being: an antisense oligonucleotide produced
CC by using (III) as a target, or a double-stranded RNA comprising a portion
CC of (III), and having the same or different sequence as that of (III); (I)
CC or (II) is useful for preparing a wound-healing agent and for treating
CC cancer. AI or EI is useful as an anti-inflammatory agent. (I) effectively
CC regulates snail activity. The present sequence represents a snail-related
CC protein of the invention. Note: This sequence is not shown in the
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences/14.07.2005/.
XX
XX Sequence 647 AA;
SQ
Query Match 5.9%; Score 7; DB 9; Length 647;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAL 24
Db 180 GGVLAAL 186
|||||
RESULT 922
ABU49815
ID ABU49815 standard; protein; 665 AA.
XX
AC ABU49815;
XX
XX 19-JUN-2003 (first entry)
DT
XX
XX Protein encoded by Prokaryotic essential gene #35342.
DB
XX
XX Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
XX Yersinia pestis.
OS
XX WO200277183-A2.
PN
XX
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002WO-US009107.
PF
XX
XX 21-MAR-2001; 2001US-00815242.
PR
XX
XX 06-SEP-2001; 2001US-00948993.
PR
XX
XX 25-OCT-2001; 2001US-0342923P.
PR

```


PR 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
PA (ELIT-) ELITRA PHARM INC.
PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyeckind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR WPI; 2003-029926/02.
DR N-PSDB; ACA53685.
XX
PT New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
PS Claim 25; SEQ ID NO 77739; 1766pp; English.
XX
CC The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 665 AA;
Query Match 5.9%; Score 7; DB 6; Length 665;
Best Local Similarity 100.0%; Pred. No. 8e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
Qy 20 VLAALAA 26
Db 399 VLAALAA 405
RESULT 923
ABG26099
ID ABG26099 standard; protein; 673 AA.
XX
AC ABG26099;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #26090.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.
XX WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS90286.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 56458; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 673 AA;
Query Match 5.9%; Score 7; DB 4; Length 673;
Best Local Similarity 100.0%; Pred. No. 8.1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;
Qy 96 QOAVIEP 102
Db 485 QOAVIEP 491
RESULT 924
ABU33621
ID ABU33621 standard; protein; 707 AA.
XX
AC ABU33621;
XX
DT 19-JUN-2003 (first entry)
XX
DE Protein encoded by prokaryotic essential gene #19148.
XX
KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
XX
KW Legionella pneumophila.
XX
PN WO200277183-A2.

```
XX 03-OCT-2002.
PD
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
PR 06-SEP-2001; 2001US-00948993.
PR
PR 25-OCT-2001; 2001US-0342923P.
PR
PR 08-FEB-2002; 2002US-00072851.
PR
PR 06-MAR-2002; 2002US-0362699P.
XX
XX (ELIT-) ELITRA PHARM INC.
PA
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
PI
XX
XX WPI; 2003-029926/02.
DR
DR N-PSDB; ACA37491.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 25; SEQ ID NO 61545; 1766pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC on a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 707 AA;
Query Match 5.9%; Score 7; DB 6; Length 707;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 GVLAALA 25
Db 640 GVLAALA 646
|||||
RESULT 925
AEB38126
ID AEB38126 standard; protein; 707 AA.
XX
XX AC AEB38126;
XX
XX 08-SEP-2005 (first entry)
L. pneumophila protein SEQ ID NO 2458.
detection; infection; Antibacterial; Vaccine.
Legionella pneumophila.
WO2005049642-A2.
PN
XX
PD 02-JUN-2005.
XX
XX 23-SEP-2004; 2004WO-IB003578.
PF
XX
XX 21-NOV-2003; 2003FR-00013687.
PR
XX
XX (INSP ) INST PASTEUR.
PA
XX (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA
XX (UPLY-) UNIV LYON 1 BERNARD CLAUDE.
PA
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Ruenick C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
PI Jarraud S;
PI
XX
XX WPI; 2005-388305/40.
DR
XX
XX New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX
XX Claim 3; SEQ ID NO 2458; 660pp; English.
PS
XX
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain; (II), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (III). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents the amino acid sequence of a
CC L. pneumophila protein.
XX
XX Sequence 707 AA;
Query Match 5.9%; Score 7; DB 9; Length 707;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 GVLAALA 25
Db 640 GVLAALA 646
|||||
RESULT 926
AEB41417
ID AEB41417 standard; protein; 707 AA.
XX
XX AC AEB41417;
XX
XX 08-SEP-2005 (first entry)
L. pneumophila protein SEQ ID NO 5749.
detection; infection; Antibacterial; Vaccine.
Legionella pneumophila.
XX
XX OS Legionella pneumophila.
```

PN WO2005049642-A2.
PD 02-JUN-2005.
XX 23-SEP-2004; 2004WO-IB003578.
XX 21-NOV-2003; 2003PR-00013687.
XX (INSP) INST PASTEUR.
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
PA (UTLY-) UNIV LYON 1 BERNARD CLAUDE.
XX (CNRS) CNRS CENT NAT RECH SCI.
PI Buchrieser C, Tichit M, Etienne J, Ma L, Cazalet C, Glaser P;
PI Rueniock C, Bouchier C, Zidane N, Magnier A, Kunst F, Vandenesch F;
XX Jarraud S;
XX WPI; 2005-388305/40.
XX New genome of Legionella pneumophila Paris strain and derived
PT polypeptides, useful for detection or identification of the strain and
PT for treatment and prevention of infections.
XX Claim 3; SEQ ID NO 5749; 660pp; English.
XX The invention relates to an isolated or purified nucleotide sequences (I)
CC from Legionella pneumophila Paris strain. (I), and their related
CC sequences or fragments, are useful as primers and probes for detection
CC and amplification, including differentiation between the Paris and
CC Philadelphia strains of Legionella pneumophila and to prepare recombinant
CC (hybrid) polypeptides (II). (II) are also useful for preparation of
CC specific antibodies (Ab), also used for detection/identification of
CC Legionella, and some (I), specifically those involved in synthesis of
CC surface proteins, are targets for identification of inhibitors. (II), or
CC vectors that contain (I), are useful as vaccines and immunogenic
CC compositions, for treatment and prevention of infections by L.
CC pneumophila. The present sequence represents the amino acid sequence of a
CC L. pneumophila protein.
XX
SQ Sequence 707 AA;
Query Match 5.9%; Score 7; DB 9; Length 707;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 19 GVLAALA 25
|||||
Db 640 GVLAALA 646
RESULT 927
ABG06939
ID ABG06939 standard; protein; 717 AA.
XX
AC ABG06939;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #6930.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
PR

XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR N-PSDB; AAS71126.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 37298; 103pp; English.
PS The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 717 AA;
Query Match 5.9%; Score 7; DB 4; Length 717;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 108 WQKLEAF 114
|||||
Db 611 WQKLEAF 617
RESULT 928
ADC87583
ID ADC87583 standard; protein; 737 AA.
XX
AC ADC87583;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human GPCR protein SEQ ID NO:2036.
XX
KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;
KW gene therapy.
XX
OS Homo sapiens.
XX
FN EP1270724-A2.
XX
PD 02-JAN-2003.
XX
XX 18-JUN-2002; 2002EP-00013517.
XX
XX 18-JUN-2001; 2001JP-00246789.
PR (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
PA

(GOLD/) GOLDMAN B S.
 Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 WPI; 2004-061375/06.
 New recombinant DNA construct comprising a promoter positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source, useful for producing plants with improved properties.
 Claim 1; SEQ ID NO 8373; 122pp; English.
 The invention relates to a recombinant DNA construct comprising a promoter functional in a plant cell, where the promoter is positioned to provide for expression of a polynucleotide encoding a polypeptide from a microbial source. The invention also relates to a transformed plant comprising the recombinant DNA construct and a method of producing a transformed plant having an improved property. The plant is a crop plant such as maize or soybean. The method of producing a transformed plant having an improved property comprises transforming a plant with the recombinant DNA construct and growing the transformed plant, where the polynucleotide or polypeptide is useful for improving plant properties. The recombinant DNA construct is useful for producing plants with improved plant properties, e.g. improved cold, heat or drought tolerance, tolerance to herbicides, extreme osmotic conditions, pathogens or pests, increased resistance to plant disease, better growth rate by modification of the cell cycle pathway with plant growth regulators, increased rate of homologous recombination, modified seed oil or protein yield and/or content, improved yield by modification of carbohydrate, nitrogen or phosphorus use and/or uptake, by modification of photosynthesis or by providing improved plant growth and development under at least one stress condition, improved lignin production or improved galactomannan production. This sequence represents a bacterial polypeptide used in the scope of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format from USPTO at seqdata.uspto.gov/sequence.html.

Query Match 5.9%; Score 7; DB 8; Length 774;
 Best Local Similarity 100.0%; Pred. No. 9.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
 |||||
 Db 335 VLALAA 341

RESULT 931
 AAG91691
 ID AAG91691 standard; protein; 785 AA.
 AC AAG91691;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum protein fragment SEQ ID NO: 5445.
 XX
 KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP1108790-A2.
 XX
 FD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 95JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX

PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A,
 XX WPI; 2001-376931/40.
 DR N-PSDB; AAH66910.
 XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 17; SEQ ID NO 5445; 246pp + Sequence Listing; English.
 XX The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium bacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
 CC corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a protein described in the
 CC exemplification of the invention. Note: The sequence data for this patent
 CC did not form part of the printed specification, but was obtained in
 CC electronic format directly from the European Patent Office
 XX
 SQ Sequence 785 AA;
 Query Match 5.9%; Score 7; DB 4; Length 785;
 Best Local Similarity 100.0%; Pred. No. 9.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ADLEVT 11
 |||||
 Db 714 ADLEVT 720

RESULT 932
 AB083221
 ID AB083221 standard; protein; 808 AA.
 XX
 AC AB083221;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Pseudomonas aeruginosa polypeptide #15396.
 XX
 KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN US6551795-B1.
 XX
 FD 22-APR-2003.
 XX
 PF 18-FEB-1999; 99US-00252991.
 XX
 PR 18-FEB-1998; 98US-0074788P.
 PR 27-JUL-1998; 98US-0094190P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX
 DR WPI; 2003-615309/58.
 DR N-PSDB; ABD16792.
 XX
 PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide, f
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX
 PS Disclosure; SEQ ID NO 31967; 455pp; English.

XX The invention relates to *Pseudomonas aeruginosa* polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a *P. aeruginosa* nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-*P. aeruginosa* drugs, as templates for recombinant
CC production of *P. aeruginosa*-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of *P. aeruginosa*-caused
CC infection, and in detection of *P. aeruginosa* sequences or other sequences
CC of *Pseudomonas* species using biochip technology. Sequences AB067826-
CC AB084396 represent *P. aeruginosa* polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
XX
SQ Sequence 808 AA;

Query Match 5.9%; Score 7; DB 7; Length 808;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 797 VLAALAA 803
|||||

RESULT 933
AAR44143
ID AAR44143 standard; protein; 816 AA.

XX AAR44143;
AC AAR44143;
DT 18-MAY-1994 (first entry)
XX Rabbit sodium ion/hydrogen ion anti-transport carrier.
DE Rabbit; sodium; hydrogen; anti-transport carrier; transgenic mouse;
XX model animal; hypertension; hypercardia.

XX Oryctolagus cuniculus.
XX JP05268856-A.
XX 19-OCT-1993.

XX 26-MAR-1992; 92JP-00068388.
XX 26-MAR-1992; 92JP-00068388.
XX (SUMU) SUMITOMO SEIYAKU KK.

XX WPI; 1993-364207/46.
XX N-PSDB; AAQ51524.

XX Transgenic mouse expressing sodium ion-hydrogen ion anti-transport
PT carrier gene - useful for model animal of hypertension and hypercardia.
XX Disclosure; Fig 8-12; 12pp; Japanese.

XX Transgenic mice are useful as animal model for the analysis of the effect
CC of sodium ion/hydrogen-ion anti-transport carrier gene activation on the
CC regulation of blood pressure. The transgenic animals are useful as models
CC of hypertension and hypercardia

XX Sequence 816 AA;

Query Match 5.9%; Score 7; DB 2; Length 816;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLGL 89
Db 738 KGVVLGL 744
|||||

RESULT 934

ABM84568
ID ABM84568 standard; protein; 816 AA.

XX ABM84568;
AC ABM84568;
DT 18-NOV-2004 (first entry)

XX Human diagnostic and therapeutic pprotein SEQ ID NO:4817.

XX gene therapy; human diagnostic and therapeutic polynucleotide; dithp.

XX Homo sapiens.

XX WO2004023973-A2.

XX 25-MAR-2004.

XX 12-SEP-2003; 2003WO-US028227.

XX 12-SEP-2002; 2002US-0410259P.

XX 12-SEP-2002; 2002US-0410260P.

XX (INCY-) INCYTE CORP.

XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F,
PI Harthorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kitton ES;
PI Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;

XX WPI; 2004-329368/30.

XX N-PSDB; ACN43220.

XX New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
FT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.

XX Claim 27; Page; 190pp; English.

XX The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm

XX Sequence 816 AA;

Query Match 5.9%; Score 7; DB 8; Length 816;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLGL 89

DB 738 KGVVLGL 744
|||||
RESULT 935
ADN26407
ID ADN26407 standard; protein; 817 AA.
XX
XX
AC ADN26407;
XX
XX
DT 02-DEC-2004 (first entry)
XX
DE Bacterial polypeptide #9060.
XX
KW Recombinant DNA construct; transformed plant; improved plant property;
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
KW pathogen tolerance; pest tolerance; plant disease resistance;
KW cell cycle pathway modification; plant growth regulator;
KW homologous recombination; seed oil yield; protein yield; carbohydrate;
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
KW bacterial polypeptide.
XX
XX Bacteria.
OS
PN US2003233675-A1.
XX
PD 18-DEC-2003.
XX
PF 20-FEB-2003; 2003US-00369493.
XX
PR 21-FEB-2002; 2002US-0360039P.
XX
PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
PI
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 9060; 122pp; English.
PS
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

SQ Sequence 817 AA;
Query Match 5.9%; Score 7; DB 8; Length 817;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 AIVPDKE 53
DB 88 AIVPDKE 94
|||||
RESULT 936
ADB99968
ID ADB99968 standard; protein; 870 AA.
XX
XX
AC ADB99968;
XX
DT 04-DEC-2003 (first entry)
XX
DE Enterohaemorrhagic E. coli O157:H7-specific protein SEQ ID NO: 14.
XX
XX enterohaemorrhagic; anti-bacterial.
OS
PN JP2002355074-A.
XX
PD 10-DEC-2002.
XX
PF 24-JAN-2002; 2002JP-00015959.
XX
PR 24-JAN-2001; 2001JP-00112010.
XX
PA (UYTS-) UNIV TSUKUBA.
XX
DR WPI; 2003-451640/43.
XX
XX Enterohaemorrhagic Escherichia coli O157:H7-specific nucleic acid molecule
PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
PI
XX Claim 3; SEQ ID NO 14; 2067pp; Japanese.
XX
XX The invention relates to a novel enterohaemorrhagic Escherichia coli
CC O157:H7-specific nucleic acid molecule. A polynucleotide of the invention
CC has anti-bacterial activity. The polypeptide can be used in detection
CC and/or treatment of O157:H7 infection. The nucleotide sequence of the
CC genome of Enterohaemorrhagic E. coli O157:H7 was determined. The present
CC sequence represents an E. coli O157:H7-specific polypeptide of the
CC invention.
SQ Sequence 870 AA;
Query Match 5.9%; Score 7; DB 7; Length 870;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALAA 26
DB 166 VLAALAA 172
|||||
RESULT 937
ADB12249
ID ADB12249 standard; protein; 879 AA.
XX
XX
AC ADB12249;
XX
DT 20-NOV-2003 (first entry)
XX
DE Alloicoccus otitis antigenic protein SEQ ID NO:6464.
XX
KW Alloicoccus otitis; antigenic protein; immunogenic; immunisation;
KW gene therapy; Gram-positive bacterium; infection.

XX OS Alloiococcus otitis.
 XX PN WO2003048304-A2.
 XX PD 12-JUN-2003.
 XX PF 25-NOV-2002; 2002WO-US036123.
 XX PR 29-NOV-2001; 2001US-0333777P.
 XX PR 18-NOV-2002; 2002US-0426742P.
 XX PA (AMHP) WYETH HOLDINGS CORP.
 XX PI Fletcher LD, McMichael JC, Russell DP, Zagursky RJ;
 XX DR WPI; 2003-505284/47.
 XX DR N-PSDB; ADB12252.
 XX PT New Alloiococcus otitis polynucleotides and polypeptides, useful for
 XX PT treating and diagnosing diseases, drug screening assays and monitoring of
 XX PT effects during drug clinical trials.
 XX PS Claim 33; SEQ ID NO 6464; 1019pp; English.
 XX CC The present invention describes an isolated polynucleotide (I) of
 CC CC Alloiococcus otitis genomic DNA, which encodes an antigenic protein.
 CC CC Alloiococcus otitis is a Gram-positive bacterium. Also described: (1)
 CC CC an isolated polypeptide that is encoded by the polynucleotide (I); (2) an
 CC CC expression vector comprising the novel isolated polynucleotide (I), its
 CC CC complement, degenerate variant or fragment; (3) a genetically engineered
 CC CC host cell, transfected, transformed or infected with the vector of (2);
 CC CC (4) an antibody specific for the polypeptide of (1); (5) an immunogenic
 CC CC composition comprising the polypeptide, its complement, biological
 CC CC equivalent or fragment, or the polynucleotide that is comprised in the
 CC CC expression vector; (6) a pharmaceutical composition comprising the
 CC CC polypeptide of (1) and a carrier; (7) a protein chip comprising an array
 CC CC of the polypeptides of (1), their biological equivalent or fragment; (8)
 CC CC immunising against Alloiococcus otitis by administering to a host the
 CC CC immunogenic composition; (9) detecting and/or identifying Alloiococcus
 CC CC otitis in the biological sample; (10) a kit comprising a container
 CC CC containing the novel polynucleotide, its degenerate variant or fragment,
 CC CC or the antibody of (4), and (11) producing a polypeptide by culturing the
 CC CC genetically engineered host cell under conditions suitable to produce the
 CC CC polypeptide from the culture. (I) can be used in gene therapy. The
 CC CC polynucleotides, polypeptides, antibodies and compositions of the present
 CC CC invention can be used for treating and diagnosing diseases, drug
 CC CC screening assays and monitoring of effects during drug clinical trials.
 CC CC The polynucleotides are useful for expressing and detecting Alloiococcus
 CC CC otitis. The present sequence represents an Alloiococcus otitis
 CC CC antigen protein from the present invention.
 XX SQ Sequence 879 AA;
 Query Match 5.9%; Score 7; DB 6; Length 879;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 84 GKVLGLL 90
 Db 448 GKVLGLL 454
 RESULT 938
 ADC00520
 ID ADC00520 standard; protein; 881 AA.
 XX AC ADC00520;
 XX DT 04-DEC-2003 (first entry)
 XX DE Enterohaemorrhagic E. coli 0157:H7-specific protein SEQ ID NO: 565.
 XX PT

KW enterohaemorrhagic; anti-bacterial.
 XX OS Escherichia coli; 0157:H7.
 XX PN JP2002355074-A.
 XX PD 10-DEC-2002.
 XX PF 24-JAN-2002; 2002JP-00015959.
 XX PR 24-JAN-2001; 2001JP-00112010.
 XX PA (UYTS-) UNIV TSUKUBA.
 XX DR WPI; 2003-451640/43.
 XX PT Enterohaemorrhagic Escherichia coli 0157:H7-specific nucleic acid molecule
 XX PT and a polypeptide and its use, a polypeptide, a vector and a host cell.
 XX PS Claim 3; SEQ ID NO 565; 2067pp; Japanese.
 XX CC The invention relates to a novel enterohaemorrhagic Escherichia coli
 CC CC 0157:H7-specific nucleic acid molecule. A polynucleotide of the invention
 CC CC has anti-bacterial activity. The polypeptide can be used in detection
 CC CC and/or treatment of 0157:H7 infection. The nucleotide sequence of the
 CC CC genome of Enterohaemorrhagic E coli 0157:H7 was determined. The present
 CC CC sequence represents an E. coli 0157:H7-specific polypeptide of the
 CC CC invention.
 XX SQ Sequence 881 AA;
 Query Match 5.9%; Score 7; DB 7; Length 881;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 20 VLAALAA 26
 Db 177 VLAALAA 183
 RESULT 939
 ABO80616
 ID ABO80616 standard; protein; 913 AA.
 XX AC ABO80616;
 XX DT 29-JUL-2004 (first entry)
 XX DE Pseudomonas aeruginosa polypeptide #12791.
 XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX OS Pseudomonas aeruginosa.
 XX PN US6551795-B1.
 XX PD 22-APR-2003.
 XX PF 18-FEB-1999; 99US-00252991.
 XX PR 18-FEB-1998; 98US-0074788P.
 XX PR 27-JUL-1998; 98US-0094190P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX DR WPI; 2003-615309/58.
 XX DR N-ESDB; ABD14187.
 XX PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 XX PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 XX PT pathological conditions resulting from bacterial infection.

XX Disclosure; SEQ ID NO 29362; 455pp; English.

XX The invention relates to Pseudomonas aeruginosa polypeptides and the

XX polynucleotides encoding them. The sequences are useful in diagnosis and

XX therapy of pathological conditions, as molecular targets for diagnostics,

XX prophylaxis and treatment of pathological conditions resulting from a

XX bacterial infection, for evaluating a compound, such as a polypeptide,

XX for the ability to bind a P. aeruginosa nucleic acid, as components of

XX effective antibacterial targets, as targets for antibacterial drugs,

XX including anti-P. aeruginosa drugs, as templates for recombinant

XX production of P. aeruginosa-derived peptides or polypeptides, as target

XX components for diagnosis and/or treatment of P. aeruginosa-caused

XX infection, and in detection of P. aeruginosa sequences or other sequences

XX of Pseudomonas species using biochip technology. Sequences ABO67826-

XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The

XX sequence data for this patent did not form part of the printed

XX specification but was obtained in electronic format from USPTO at

XX seqdata.uspto.gov/sequence.html

XX

XX Sequence 913 AA;

Query Match 5.9%; Score 7; DB 7; Length 913;

Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26

Db 198 VLAALAA 204

|||||

RESULT 940

ADB64658

ID ADB64658 standard; protein; 1036 AA.

AC ADB64658;

XX

XX 04-DEC-2003 (first entry)

DE Human protein encoded by clone MESAN20027900.

DE

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;

KW cell regeneration; membrane protein; signal transduction-related protein;

KW transcription-related protein; osteoporosis; neurological disease;

KW cancer; tumour.

XX

OS Homo sapiens.

XX

PN EP1308459-A2.

XX

XX 07-MAY-2003.

XX

XX 28-MAR-2002; 2002EP-00007401.

PF

XX

XX 05-NOV-2001; 2001JP-00379298.

PR

XX

XX 25-JAN-2002; 2002US-00350978.

PR

XX

XX (HELI-) HELIX RES INST.

PA

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;

XX

XX WPI; 2003-450961/43.

DR

XX N-PSDB; ADB62688.

DR

XX

XX New polynucleotides and polypeptides, useful for developing a diagnostic

PT marker or medicines for regulation of their expression and activity, or

PT as targets of gene therapy.

XX

XX Claim 1; Page; 222pp; English.

PS

XX

CC The invention discloses a polynucleotide comprising a sequence selected

CC from 1970 fully defined nucleotide sequences which encode novel

CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide

CC or its partial peptide, an antibody binding to the polypeptide or peptide

CC of the polynucleotide, immunologically assaying the polypeptide or

CC peptide of the polynucleotide by contacting the polypeptide or peptide

CC with the antibody of the encoded protein, and observing the binding

CC between the two, a transformant carrying the polynucleotide in an

CC expressible manner and an antisense polynucleotide. The oligonucleotide

CC is useful as a primer for synthesising the polynucleotide, or as a probe

CC for detecting the polynucleotide. The polynucleotides and encoded

CC proteins are useful as pharmaceutical agents and many disease-related

CC genes may be included in them, for developing a diagnostic marker or

CC medicines for regulation of their expression and activity, or as targets

CC of gene therapy. The genes are involved in tissue and/or cell

CC regeneration. Membrane proteins, signal transduction-related proteins,

CC transcription-related proteins, disease-related proteins and genes

CC encoding them can be used as indicators for diseases (e.g. osteoporosis,

CC neurological diseases, cancer, tumours. The cDNA may be used to regulate

CC the activity or expression of the encoded protein to treat diseases. The

CC sequence presented is a protein of the invention. Note: Some of the

CC sequence data for this patent is not represented in the printed

CC specification, but is based on sequence information supplied by the

CC European Patent Office.

XX

XX Sequence 1036 AA;

Query Match 5.9%; Score 7; DB 7; Length 1036;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 QQQAVIE 101

Db 27 QQQAVIE 33

|||||

RESULT 941

ABU21740

ID ABU21740 standard; protein; 1082 AA.

XX

XX ABU21740;

DT 19-JUN-2003 (first entry)

XX

XX Protein encoded by Prokaryotic essential gene #7267.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Burkholderia fungorum.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA25610.

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 49664; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

XX the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 1082 AA;

Query Match 5.9%; Score 7; DB 6; Length 1082;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLA 22

Db 601 LGGVLA 607

|||||

RESULT 942

ADF04094

ID ADF04094 standard; protein; 1116 AA.

XX

AC ADF04094;

XX

DT 12-FEB-2004 (first entry)

XX

DE Bacterial polypeptide #207.

XX

KW Proteus mirabilis infection; bacterial infection; antibacterial;

KW immunostimulant.

XX

OS Proteus mirabilis.

XX

PN US6605709-B1.

XX

PD 12-AUG-2003.

XX

PF 05-APR-2000; 2000US-00543681.

XX

PF 09-APR-1999; 99US-0128706P.

XX

XX (GENO-) GENOME THERAPEUTICS CORP.

PA Breton GL;

XX

PI

XX

PT

XX

PS WPI; 2003-895291/82.

DR N-PSDB; ADE99921.

XX

XX New Proteus mirabilis polypeptides and polynucleotides, useful as

PT reagents for diagnosis of bacterial disease, as components of

PT antibacterial vaccines, as targets for antibacterial drugs, or as

PT biocontrol agents for plants.

XX

PS Disclosure; SEQ ID NO 4379; 870pp; English.

XX

XX The invention relates to new Proteus mirabilis polypeptides and

CC polynucleotides. The invention also relates to antibodies against the

CC polypeptides, methods for producing the polypeptides, a method of

CC generating vaccines for immunising an individual against *P. mirabilis*, a

CC method for evaluating a compound for the ability to bind a *P. mirabilis*

CC polypeptide and a method for screening test compounds for anti-bacterial

CC activity. The polypeptides and polynucleotides are useful as molecular

CC targets for diagnosing, preventing and treating pathological conditions

CC resulting from bacterial infection, as reagents for diagnosis of

CC bacterial diseases, as components of antibacterial vaccines, as targets

CC for antibacterial drugs or as bio-control agents for plants. This

CC sequence represents a Proteus mirabilis polypeptide of the invention.

XX

SQ Sequence 1116 AA;

Query Match 5.9%; Score 7; DB 7; Length 1116;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26

Db 689 VLAALAA 695

|||||

RESULT 943

ADJ87387

ID ADJ87387 standard; protein; 1255 AA.

XX

AC ADJ87387;

XX

DT 06-MAY-2004 (first entry)

XX

DE DNA replication and pathogenesis associated protein RXA01770.

XX

KW DNA-replication protein; ribosome protein; pathogenesis protein;

KW amino acid production; lysine production; food; animal feed; cosmetic;

KW pharmaceutical.

XX

OS Corynebacterium glutamicum.

XX

FH Key Location/Qualifiers

FT Misc-difference 243

FT /label= D243N

XX

PN DE10154246-A1.

XX

PD 08-MAY-2003.

XX

PF 05-NOV-2001; 2001DE-01054246.

XX

PF 05-NOV-2001; 2001DE-01054246.

XX

XX (BADI) BASF AG.

XX

XX Zeider O, Pompejus M, Schroeder H, Kroeger B, Klopprogge C;

PI Haberhauer G;

XX

XX WPI; 2003-431903/41.

DR N-PSDB; ADJ87386.

XX

XX New nucleic acid encoding variant forms of DNA-replication, ribosome and

PT pathogenesis proteins, useful for production of fine chemicals,

PT specifically lysine, in microorganisms.

XX Claim 1; SEQ ID NO 44; 259pp; German.
 PS This invention describes novel DNA-replication, ribosome and pathogenesis
 CC proteins from *Corynebacterium glutamicum*, involved directly or indirectly
 CC in formation of fine chemicals. The proteins are isolated from a nucleic
 CC acid library of *C. glutamicum* then mutated at specified positions, cloned
 CC and expressed by standard methods. Cells containing vectors that express
 CC the proteins are used for production of fine chemicals, preferably amino
 CC acids and specifically lysine, but more generally nucleotides,
 CC nucleosides, lipids, fatty acids, diols, carbohydrates, aromatic
 CC compounds, vitamins, co-factors and enzymes. These are useful in the
 CC food, animal feed, cosmetics and pharmaceutical industries. The encoding
 CC polynucleotides optionally as primers and probes, can also be used for
 CC identification and classification of *C. glutamicum* and related species,
 CC e.g. for diagnosis, for genomic mapping, functional or evolutionary
 CC studies, gene manipulation and modulation of metabolic activity. ADJ87344
 CC -ADJ87439 represent *C. glutamicum* DNA-replication, ribosome and
 CC pathogenesis-associated polynucleotides and polypeptides described in the
 CC disclosure of the invention.
 XX
 SQ Sequence 1255 AA;

Query Match 5.9%; Score 7; DB 7; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 11
 |||||
 DB 1184 ADLEVT 1190

RESULT 944
 ADL65837
 ID ADL65837 standard; protein; 1255 AA.
 XX
 AC ADL65837;
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE *C. glutamicum* RXA-associated protein #97.
 XX
 KW fine chemical production; lysine production; nucleoside;
 KW lipid; fatty acid; diol; carbohydrate; aromatic compound; vitamin;
 KW co-factor; enzyme; food; animal feed; cosmetic; pharmaceutical.
 XX
 OS *Corynebacterium glutamicum*.
 XX
 PN DE10154177-A1.
 XX
 PD 08-MAY-2003.
 XX
 PF 05-NOV-2001; 2001DE-01054177.
 XX
 PR 05-NOV-2001; 2001DE-01054177.
 XX
 PA (BADI) BASF AG.
 XX
 PI Zelder O, Pompejus M, Schroeder H, Kroeger B, Klopprogge C;
 PI Haberhauer G;
 XX
 DR WPI; 2003-431900/41.
 DR N-PSDB; ADL65836.
 XX
 PT New nucleic acid encoding variant forms of marker and fine chemical-
 PT production proteins, useful for production of fine chemicals,
 PT specifically lysine, in microorganisms.
 XX
 PS Claim 1; Page; 20pp; German.
 XX
 CC This invention describes novel polynucleotides that encode protein
 CC markers and fine chemical-production proteins from *Corynebacterium*
 CC glutamicum. The polynucleotides are isolated from a nucleic acid library

CC of *C. glutamicum* then mutated at the specified positions, cloned and
 CC expressed by standard methods. Cells, especially *Corynebacterium*
 CC glutamicum, containing vectors that express the polynucleotides are used
 CC for production of fine chemicals, preferably amino acids and specifically
 CC lysine, but more generally nucleotides, nucleosides, lipids, fatty acids,
 CC diols, carbohydrates, aromatic compounds, vitamins, co-factors and
 CC enzymes. These are useful in the food, animal feed, cosmetics and
 CC pharmaceutical industries. The polynucleotides, optionally as primers and
 CC probes, can also be used for identification and classification of *C.*
 CC glutamicum and related species, e.g. for diagnosis, for genomic mapping,
 CC functional or evolutionary studies, gene manipulation and modulation of
 CC metabolic activity. Cells that containing the polynucleotides of the
 CC invention may produce fine chemicals in better yields, with higher
 CC productivity and/or more efficiently. NOTE: This sequence is not
 CC represented in the printed specification but is available in electronic
 CC format. The sequence represented in this record has been obtained from
 CC WO2003046123.
 XX

SQ Sequence 1255 AA;

Query Match 5.9%; Score 7; DB 7; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 11
 |||||
 DB 1184 ADLEVT 1190

RESULT 945
 AAU84801
 ID AAU84801 standard; protein; 2010 AA.
 XX
 AC AAU84801;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE HCV HepC cassette B.
 XX
 KW Savine; vaccine; cancer; viral infection; HIV; hepatitis C virus;
 KW viral infection; human immunodeficiency virus; melanoma;
 KW bacterial infection; Salmonella; Legionella; parasitic infection;
 KW Trypanosoma; Toxoplasma; Giardia.
 XX
 OS Hepatitis C virus.
 XX
 PN WO200190197-A1.
 XX
 PD 29-NOV-2001.
 XX
 PF 25-MAY-2001; 2001WO-AU000622.
 XX
 PR 26-MAY-2000; 2000AU-00007761.
 XX
 PA (AUSU) UNIV AUSTRALIAN NAT.
 XX
 PI Thomson SA, Ramshaw IA;
 XX
 DR WPI; 2002-147575/19.
 DR N-PSDB; ABK36639.
 XX
 PT New synthetic polypeptides having several different segments of at least
 PT one parent polypeptide linked together differently compared to the
 PT linkage in the parent polypeptide, for inducing immune response against a
 PT pathogen or cancer.
 XX
 PS Example 2; SEQ ID NO 814; 364pp; English.
 XX
 CC The invention relates to a new synthetic polypeptide (I) comprising
 CC several different segments of at least one parent polypeptide linked
 CC together in a different relationship relative to their linkage in the
 CC parent polypeptide to impede, abrogate or otherwise alter at least one
 CC function associated with the parent polypeptide and for inducing an

CC immune response against a pathogen or cancer. Also included are a
 CC synthetic polynucleotide encoding and a computer system for designing the
 CC synthetic polypeptides. The synthetic polypeptides and polynucleotides
 CC are referred to as a vaccine. The synthetic polypeptide is useful for
 CC modulating immune responses preferably directed against a pathogen or a
 CC cancer, (e.g., cancers of the lung, breast, ovary, cervix, colon, head
 CC and neck, pancreas, prostate, stomach, bladder, kidney, bone liver,
 CC oesophagus, brain, testicle, uterus), as potentiating agents.
 CC Compositions comprising the polypeptide may be used in the treatment or
 CC prophylaxis against viral (such as infections caused by HIV (human
 CC immunodeficiency virus), hepatitis, influenza, Japanese encephalitis
 CC virus, Epstein-Barr virus and respiratory syncytial virus), bacterial
 CC (e.g., infections caused by Neisseria, Meningococcal, Haemophilus,
 CC Salmonella, Streptococcal, Legionella and Mycobacterium) or parasitic
 CC (e.g., infections caused by Plasmodium, Schistosoma, Leishmania,
 CC Trypanosoma, Toxoplasma and Giardia) infections. The present sequence is
 CC a cassette protein consisting of several peptides derived from a parent
 CC protein. One or more cassettes are used to construct a vaccine of the
 CC invention

XX SQ Sequence 2010 AA;

Query Match 5.9%; Score 7; DB 5; Length 2010;
 Best Local Similarity 100.0%; Pred. No. 2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNWQKLE 112
 Db 1199 TNWQKLE 1205
 |||||

RESULT 946

AA53677
 ID AAY53677 standard; protein; 2278 AA.

AC AAY53677;

DT 22-FEB-2000 (first entry)

DE Sequence gi/3413886/dbj/BAA323071 from an alignment with protein 274.

KW Mechanical stress; gene therapy; protein 274; osteoporosis; bone density;
 KW bone development; gi/328186.

XX Unidentified.

XX WO9960164-A1.

XX 25-NOV-1999.

XX 14-MAY-1999; 99WO-US011066.

XX 15-MAY-1998; 98US-0085673P.

XX (QUAR-) QUARK BIOTECH INC.

XX Einat P, Mor O, Skalliter R, Feinstein E, Faerman A;

XX WPI; 2000-053304/04.

PT Identification of stress induced genes for determining risk and
 PT preventing, treating or controlling osteoporosis.

XX Claim 32; Fig 14A-T; 308pp; English.

XX The present sequence is obtained from a clustal X alignment with protein
 CC 274. Protein 274 was identified using the method of the invention after
 CC subjecting rat osteoblasts to mechanical stress. Expression of the 608
 CC gene was found to be upregulated by about 3-fold in cells subjected to
 CC mechanical strain. The specification describes a method for the
 CC identification of genes responsive to a specific mechanical stress. The
 CC method comprises applying the mechanical stress to an organism (tissue or
 CC cells comprising bone cells), isolating the specific cellular fractions

CC and extracting mRNA from them, and differentially analysing the mRNA in
 CC comparison with control samples. The method is used to identify genes
 CC whose expression is responsive to a specific stress. The identified genes
 CC are employed in determining risk associated with a physiological or
 CC disease state. The risk determination methods are used for testing a
 CC medicament for gene therapy. These medicaments, or genes identified by
 CC the method of the invention, are used for treating, preventing or
 CC controlling a physiological or disease state (especially osteoporosis or
 CC bone density or other factors causing or contributing to osteoporosis or
 CC its symptoms or other conditions involved in mechanical stress or its
 CC lack. The methods can also be used for advancing research or studies in
 CC bone development

XX SQ Sequence 2278 AA;

Query Match 5.9%; Score 7; DB 3; Length 2278;
 Best Local Similarity 100.0%; Pred. No. 2.2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 Db 425 VLAALAA 431
 |||||

RESULT 947

ABU70437
 ID ABU70437 standard; protein; 2408 AA.

AC ABU70437;

DT 10-JUN-2003 (first entry)

DE Human adipocyte Selected Interacting domain, SID, #68.

XX Human; prey; adipocyte; SID; selected interacting domain; anorectic;
 KW anti-diabetic; protein-protein interaction; diabetes;
 KW yeast 2-hybrid assay; metabolic disorder; obesity.

XX Homo sapiens.

XX WO200286122-A2.

XX 31-OCT-2002.

XX 14-MAR-2002; 2002WO-EP003768.

XX 14-MAR-2001; 2001US-0275734P.

XX (HYBR-) HYBRIGENICS.

XX Legrain P, Daviet L;

XX WPI; 2003-103412/09.

XX N-PSDB; ACA56981.

PT New complex between two interacting proteins in adipocyte cells, useful
 PT for identifying selected interacting domains that modulate protein
 PT interactions, or for preventing or treating metabolic disorders such as
 PT obesity or diabetes.

XX Claim 6; Page 134-137; 382pp; English.

XX The invention relates to a complex between two interacting proteins in
 CC adipocyte cells, given in the specification. The proteins are identified
 CC by selecting a bait protein from a known adipocyte marker and then
 CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
 CC members of an adipocyte cDNA library. The proteins are designated as a
 CC (RIM) (selected interacting domains) proteins. Also included are a
 CC polynucleotide encoding a polypeptide in the adipocyte cells, a
 CC recombinant host cell expressing at least one of the interacting
 CC polypeptides of the complex, selecting a modulating compound in adipocyte
 CC cells, a SID (RIM) polypeptide comprising any of the 738 amino acid
 CC sequences given in the specification (including its fragment or variant),

XX PS Example 4; SEQ ID NO 257; 107pp; English.

XX CC The invention relates to novel isolated human secreted polypeptides (I) and polynucleotides (II). (I) and (II) are useful for treating inflammatory conditions such as arthritis, nephritis, Crohn's disease, ischaemia-reperfusion injury, shock, sepsis, immune responses, and is involved in increasing haematopoiesis, stem cell survival, bone growth and remodeling. (I), (II) and modulators of (II) are useful for prophylaxis or treatment of one or more cancers. (II) is also useful for creating transgenic animals useful for studying the in vivo activities of the polypeptide as well as for studying modulators of the polypeptides. (I) induces the proliferation of neural cells and regeneration of nerve and brain tissue and is useful for the treatment of central and peripheral nervous system diseases and neuropathies, such as Alzheimer's, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis. In addition, (I) is involved in chemotactic or chemokinetic activity, regulation of haematopoiesis and is useful for treating myeloid or lymphoid cell disorders, platelet disorders such as thrombocytopenia and for regeneration of bone, cartilage, tendon, ligament and/or nerve tissue growth, and in tissue repair, healing of burns, incisions, ulcers, for treating osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease. Furthermore, (I) is also useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, various immune deficiencies and disorders including severe combined immunodeficiency (SCID), bacterial or fungal infections, autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis, diabetes mellitus, myasthenia gravis, allergic reactions and conditions, such as asthma or other respiratory problems. In addition, (I) affects biorythms or circadian cycles of rhythms, fertility, metabolism, catabolism, anabolism, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, provides analgesic effects or other pain reducing effects, immunoglobulin like activity and can act as an antigen in a vaccine composition to raise an immune response. AAU28020-AAU28395 represent novel human secreted protein amino acid sequences of the invention

XX SQ Sequence 2458 AA;

Query Match 5.9%; Score 7; DB 4; Length 2458;
 Best Local Similarity 100.0%; Pred. No. 2.4e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAALAA 26
 |||||
 Db 607 VLAAALAA 613

RESULT 950
 ADC42917
 ID ADC42917 standard; protein; 2940 AA.

XX AC ADC42917;
 XX CC
 XX DT 18-DEC-2003 (first entry)
 XX DE Hepatitis C virus protein.
 XX KW viral; YPXL; cellular protein AP-50; hepatitis C virus; HCV;
 KW herpes simplex virus; HSV; UL42 protein; variola virus A10L;
 KW vaccinia virus virion core protein Pxa;
 KW human parainfluenza virus haemagglutinin-neuraminidase; HSP-1;
 KW viral budding; viral infectivity.
 XX OS Hepatitis C virus.
 XX PN WO2003017943-A2.
 XX PD 06-MAR-2003.
 XX PR 22-AUG-2002; 2002WO-US027066.
 XX PR 22-AUG-2001; 2001US-0314182P.

XX PA (MYRI-) MYRIAD GENETICS INC.
 XX PI Morham S, Zavitz K, Hobden A;
 XX DR WPI; 2003-371696/35.
 XX PT Novel peptide, comprising a contiguous amino acid sequence of a viral protein, capable of binding to a region of cellular protein AP-50 is useful for treating viral infections.
 XX PS Example 3; SEQ ID NO 13; 60pp; English.
 XX CC The invention relates to a novel isolated peptide comprising a contiguous amino acid sequence of 7-30 amino acid residues of a viral protein, where the contiguous amino acid sequence encompasses the YPXL motif of the viral protein, and the 7-30 amino acid viral protein is capable of binding a region including the amino acid residues 121-435 of cellular protein AP-50. The isolated 7-30 amino acid viral protein is selected from Hepatitis C virus (HCV) polyprotein, herpes simplex virus (HSV) UL42 protein, variola virus A10L protein, vaccinia virus virion core protein Pxa, and human parainfluenza virus haemagglutinin-neuraminidase, where the contiguous amino acid sequence encompasses the YPXL motif of the viral protein. The peptides of the invention are useful in the manufacture of a medicament for treating viral infections caused by a virus selected from HCV, HSP-1 and variola virus. The peptides are also useful for inhibiting viral budding from virus-infected cells and/or inhibiting viral infectivity, thus inhibiting viral propagation in the cells. This sequence represents the protein of a Hepatitis C virus used in the treatment of viral infections of the invention.

XX SQ Sequence 2940 AA;

Query Match 5.9%; Score 7; DB 7; Length 2940;
 Best Local Similarity 100.0%; Pred. No. 2.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLV 56
 |||||
 Db 1700 PDKEVLV 1706

RESULT 951
 ADE40130
 ID ADE40130 standard; protein; 3006 AA.

XX AC ADE40130;
 XX DT 29-JAN-2004 (first entry)
 XX DE Human NOV15a protein - SEQ ID 36.
 XX KW NOVX; cardiac; antiarteriosclerotic; hypotensive; cytostatic; anorectic; antidiabetic; immunosuppressive; anti-HIV; neuroprotective; nootropic; antiparkinsonian; antisthmatic; gynaecological; cardiomyopathy; atherosclerosis; hypertension; cancer; obesity; diabetes; AIDS; multiple sclerosis; graft-versus-host disease; Alzheimer's; Parkinson's; asthma; fertility disorder; vaccine; gene therapy; chromosome mapping; tissue typing; human; NOV.
 XX OS Homo sapiens.
 XX PN WO2003064589-A2.
 XX PD 07-AUG-2003.
 XX PF 02-AUG-2002; 2002WO-US024483.
 XX PR 02-AUG-2001; 2001US-0309501P.
 PR 03-AUG-2001; 2001US-0310291P.
 PR 07-AUG-2001; 2001US-0310544P.
 PR 08-AUG-2001; 2001US-0310951P.
 PR 09-AUG-2001; 2001US-0311292P.

PR 13-AUG-2001; 2001US-0311979P.
 PR 16-AUG-2001; 2001US-0312892P.
 PR 17-AUG-2001; 2001US-0313201P.
 PR 17-AUG-2001; 2001US-0313415P.
 PR 20-AUG-2001; 2001US-0313643P.
 PR 20-AUG-2001; 2001US-0313702P.
 PR 21-AUG-2001; 2001US-0314031P.
 PR 23-AUG-2001; 2001US-0314466P.
 PR 28-AUG-2001; 2001US-0315403P.
 PR 29-AUG-2001; 2001US-0315853P.
 PR 17-SEP-2001; 2001US-0322716P.
 PR 21-SEP-2001; 2001US-0323994P.
 PR 14-DEC-2001; 2001US-0340233P.
 PR 05-FEB-2002; 2002US-0354591P.
 PR 19-MAR-2002; 2002US-0365478P.
 PR 19-APR-2002; 2002US-0373814P.
 PR 19-APR-2002; 2002US-0373825P.
 PR 19-APR-2002; 2002US-0373989P.
 PR 23-APR-2002; 2002US-0374632P.
 PR 07-JUN-2002; 2002US-0386971P.
 PR 01-AUG-2002; 2002US-00210172.
 PA (CURA-) CURAGEN CORP.
 XX
 XX
 PI Kekuda R, Miller CE, Patturajan M, Pena CEA, Rieger DK;
 PI Shinkets RA, Zerkusen BD, Li L, Ji W, Padigar M, Casman SJ;
 PI Voss EZ, Boldog FL, Gorman L, Leite MW, Vernet CM, Anderson DW;
 PI Guo X, Zhong M, Garlach VL, Hjalt T, Rastelli L, Spytek KA;
 PI Edinger SR, Ellerman K, Malvankar UM, Macdougall JR, Stone DJ;
 PI Alsebrook JP, Lepley DM, Burgess CE, Majumder K, Wolenc AR;
 PI Smithson G;
 XX
 DR WPI; 2003-663472/62.
 DR N-PSDB; ADE40129.
 XX
 XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 1; SEQ ID NO 16; 560pp; English.
 XX
 CC The invention relates to a novel NOVX polypeptide. The polypeptide of the
 CC invention demonstrates cardiac, antiarteriosclerotic, hypotensive,
 CC cytosstatic, anorectic, antidiabetic, immunosuppressive, anti-HIV,
 CC neuroprotective, nootropic, antiparkinsonian, antiasthmatic and
 CC synaenological activities and may be useful in diagnosing, treating or
 CC preventing NOVX-associated disorders including cardiomyopathy,
 CC atherosclerosis, hypertension, cancer, obesity, diabetes, AIDS, multiple
 CC sclerosis, graft-versus-host disease, Alzheimer's disease, Parkinson's
 CC disease, asthma or fertility disorders. Furthermore, the polypeptides may
 CC be utilised as vaccines whilst the nucleic acids may be used as
 CC hybridisation probes, in gene therapy, chromosome mapping, tissue typing,
 CC preventive medicine and pharmacogenomics. The current sequence is that of
 CC the human NOV protein of the invention.
 XX
 SQ Sequence 3006 AA;
 Query Match 5.9%; Score 7; DB 7; Length 3006;
 Best Local Similarity 100.0%; Pred. No. 2.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLAALAA 26
 |||||
 Db 1155 VLAALAA 1161
 RESULT 952
 AAR33538
 ID AAR33538 standard; protein; 3033 AA.
 XX
 AC AAR33538;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 XX 01-JUL-1993 (first entry)
 DE NANBH virus strain HC-J6 protein.
 XX
 KW Non A non B hepatitis virus; plasma.
 XX Non-A.
 OS non-B hepatitis virus.
 XX
 PN EP532167-A2.
 XX
 PD 17-MAR-1993.
 XX
 PF 30-JUL-1992; 92EP-00306952.
 XX
 PR 09-AUG-1991; 91JP-00287402.
 PR 05-DEC-1991; 91JP-00360441.
 XX
 PA (IMMO) IMMUNO JAPAN INC.
 XX
 PI Okamoto H, Nakamura T;
 XX
 DR WPI; 1993-087166/11.
 DR P-PSDB; AAQ38218.
 XX
 PT Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus - useful
 PT for detecting NANBH, as a vaccine and for screening blood samples.
 XX
 PS Claim 5; Page 38-52; 93pp; English.
 XX
 CC RNA was isolated from the plasma of human patients positive for NANBH
 CC virus (strain HC-J6) and was subjected to reverse transcription to
 CC produce cDNA. The resulting cDNA was amplified by PCR, and nucleic acid
 CC sequences determined by analysis of both clones from the cDNA library and
 CC clones obd. by PCR amplification (36 clones in total). The NANBH HC-J6
 CC genome was found to contain an open reading frame encoding a polypeptide
 CC precursor of 3033 amino acid residues. See also AAR33539 and AAR33514.
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX
 SQ Sequence 3033 AA;
 Query Match 5.9%; Score 7; DB 2; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 2.9e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 50 PDKEVLY 56
 |||||
 Db 1700 PDKEVLY 1706
 RESULT 953
 AAB59172
 ID AAB59172 standard; protein; 3033 AA.
 XX
 AC AAB59172;
 XX
 DT 21-MAR-2001 (first entry)
 XX
 DE Protein encoded by infectious Hepatitis C virus 2a genotype.
 XX
 KW GBV-B; hepatitis C virus; HCV; vaccine.
 XX
 OS Hepatitis C virus.
 XX
 PN WO200075337-A1.
 XX
 PD 14-DEC-2000.
 XX
 PF 02-JUN-2000; 2000WO-US015293.
 XX

```

PR 04-JUN-1999; 99US-0137694P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Bukh J, Yanagi M, Emerson SU, Purcell RH;
XX WPI; 2001-091214/10.
XX New infectious nucleic acids of the GB virus-B clone, useful for
PT indirectly studying the molecular properties of hepatitis C virus (HCV)
PT and in developing vaccines and therapeutics for HCV.
XX Disclosure; Page 82-94; 96pp; English.
XX The present invention relates to GB virus-B. The nucleic acid molecules
CC of the invention are useful for indirectly studying the molecular
CC properties of hepatitis C virus (HCV). The infectious nucleic acid
CC sequence of the GB virus-B clone and the HCV/GBV-B chimeras may be used
CC in the development of vaccines and therapeutics for HCV
XX Sequence 3033 AA;
SQ
Query Match 5.9%; Score 7; DB 4; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
DB 1700 PDKEVLY 1706
|||||

RESULT 954
AAB31168
ID AAB31168 standard; protein; 3033 AA.
XX AAB31168;
AC AAB31168;
XX 02-APR-2001 (first entry)
DT 02-APR-2001 (first entry)
XX Amino acid sequence of a hepatitis C virus (HCV) clone genotype 2a.
DE Chimeric virus; bovine viral diarrhoea virus; BVDV; hepatitis C virus;
KW HCV; vaccine; viral inhibitor; antiviral.
XX Hepatitis C virus.
OS WO200075352-A2.
XX 14-DEC-2000.
PD 14-DEC-2000.
XX 02-JUN-2000; 2000WO-US015527.
PF 04-JUN-1999; 99US-0137817P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Nam J, Bukh J, Emerson SU, Purcell RH;
XX WPI; 2001-071081/08.
XX N-PSDB; AAC86937.
XX New nucleic acid comprising a chimeric bovine viral diarrhoea virus genome
PT in which the (non-)structural region has been replaced by hepatitis C
PT virus (HCV) genome useful for treating or preventing HCV signs and
PT symptoms.
XX Disclosure; Page 85-97; 97pp; English.
XX The specification describes a nucleic acid comprising a chimeric virus
CC genome, specifically bovine viral diarrhoea virus (BVDV) genome in which
CC the (non-)structural region has been replaced by the (non-)structural
CC region of a hepatitis C virus (HCV) genome. The nucleic acids comprising
CC the chimeric virus and the chimeric virus are useful for identifying cell
CC
lines capable of supporting the replication of these chimeric viruses, in
CC screening for neutralizing antibodies to HCV of different genotypes, in
CC the production of HCV-BVDV virions, for the development of inactivated or
CC attenuated vaccines to prevent HCV-BVDV in a mammal, in studying the
CC molecular properties of HCV indirectly in vitro, and in identifying
CC inhibitors of viral enzyme activity which would be useful as antiviral
CC agents. Formulations or compositions comprising the chimeric virions may
CC be used to treat or prevent the signs and symptoms of HCV. The present
CC sequence is encoded by a HCV clone, which is used to construct chimeric
CC nucleic acids of the invention
XX Sequence 3033 AA;
SQ
Query Match 5.9%; Score 7; DB 4; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
DB 1700 PDKEVLY 1706
|||||

RESULT 955
AAB30729
ID AAB30729 standard; protein; 3033 AA.
XX AAB30729;
AC AAB30729;
XX 02-APR-2001 (first entry)
DT 02-APR-2001 (first entry)
XX Amino acid sequence of infectious Hepatitis C virus strain HC-J6CH.
DE HCV; HCV strain HC-J6CH; HCV genotype 2a; antiviral; vaccine.
KW Hepatitis C virus.
OS WO200075338-A2.
XX 14-DEC-2000.
PD 14-DEC-2000.
XX 02-JUN-2000; 2000WO-US015446.
PF 04-JUN-1999; 99US-0137693P.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA Yanagi M, Bukh J, Emerson SU, Purcell RH;
XX WPI; 2001-061728/07.
XX N-PSDB; AAC86644.
XX Nucleic acid molecule encoding human hepatitis C virus of genotype 2a for
PT developing vaccines, for diagnosis of hepatitis C virus and in screening
PT assays for identification of antiviral agents.
XX Disclosure; Page 88-99; 167pp; English.
XX The present sequence represents an amino acid sequence of infectious
CC Hepatitis C virus (HCV) strain HC-J6CH genotype 2a. The HCV
CC polynucleotide sequence is capable of expressing the virus when
CC transfected into cells. The HCV protein is useful for assaying candidate
CC antiviral agents for activity against HCV. Antibodies specific for HCV
CC polypeptide are useful in prevention and treatment of diseases caused by
CC HCV in animals, in particular humans. The HCV polypeptides serve as
CC immunogens in the development of vaccines for preventing HCV in mammals
CC or as antigens in diagnostic assays for detecting the presence of HCV in
CC biological samples. The HCV polynucleotide is also useful for identifying
CC cell lines capable of supporting the replication of HCV in vitro and to
CC produce attenuated viral strains via passage in vitro or in vivo
XX Sequence 3033 AA;
SQ
Query Match 5.9%; Score 7; DB 4; Length 3033;

```


Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 956
ABG30688
ID ABG30688 standard; protein; 3033 AA.
AC ABG30688;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human HCV-related polypeptide.
XX
KW Hepatitis C virus; human; virucide; gene therapy; HCV;
KW fulminant hepatitis C.
XX
OS Homo sapiens.
XX
FN JP2002171978-A.
XX
PD 18-JUN-2002.
XX
PF 01-DEC-2000; 2000JP-00367365.
XX
PR 01-DEC-2000; 2000JP-00367365.
XX
PA (TOKR-) ZH TOKYO RINSHO IGAKU SOGO KENKYUSHO.
PA (TORA) TORAY IND INC.
XX
DR WPI; 2002-569884/61.
DR N-PSDB; ABK88904.
XX
XX A gene of a fulminant hepatitis C virus strain and the encoded
PT polypeptide useful in gene therapy to treat hepatitis C.
XX
PS Claim 1; Page 25-33; 36pp; Japanese.
XX
CC The invention relates to a human polypeptide related to hepatitis C virus
CC (HCV), and the polynucleotide encoding it. The polypeptide can be used
CC for the development of gene therapy on fulminant hepatitis C. This
CC sequence represents a human HCV-related polypeptide
XX
SQ Sequence 3033 AA;

Query Match 5.9%; Score 7; DB 5; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 957
ADV04740
ID ADV04740 standard; protein; 3033 AA.
XX
AC ADV04740;
XX
DT 24-FEB-2005 (first entry)
XX
DE Hepatitis C virus (HCV) protein #2.
XX
KW Virucide; hepatitis C virus infection; replicon.
XX
OS Hepatitis C virus.
XX
FN WO2004104198-A1.

Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 958
ADV04738
ID ADV04738 standard; protein; 3033 AA.
XX
AC ADV04738;
XX
DT 24-FEB-2005 (first entry)
XX
DE Hepatitis C virus (HCV) protein #1.
XX
KW Virucide; hepatitis C virus infection; replicon.
XX
OS Hepatitis C virus.
XX
FN WO2004104198-A1.
XX
PD 02-DEC-2004.
XX
PF 25-NOV-2003; 2003WO-JP015038.

XX 02-DEC-2004.
XX 25-NOV-2003; 2003WO-JP015038.
XX 26-MAY-2003; 2003JP-00148242.
XX 19-SEP-2003; 2003JP-00329115.
XX (TORA) TORAY IND INC.
XX (TOKM-) TOKYO METROPOLITAN ORG MEDICAL RES.
XX (UTMA-) UNIV MAINZ GUTENBERG JOHANNES.
XX Wakita T, Kato T, Date T;
XX WPI; 2005-013292/01.
XX N-PSDB; ADV04739.
XX Novel replicon RNA, having sequence of 5' and-3' untranslated region and
XX base sequence encoding NS3, NS4A, NS4B, NS5A and NS5B proteins on genomic
XX RNA of hepatitis C virus of genotype 2a, useful for treating hepatitis C
XX virus infection.
XX Disclosure; SEQ ID NO 6; 197pp; Japanese.
XX The invention relates to replicon RNA from genotype 2a of hepatitis C
XX virus comprising a 5' untranslated region, a base sequence encoding NS3
XX protein, NS4A protein, NS4B protein, NS5A protein and NS5B protein, and a
XX 3' untranslated region. The invention also relates to a cell capable of
XX reproducing the replicon involving transducing the replicon RNA to a
XX cell, a method of producing a hepatitis C virus protein, a method of
XX screening a substance that promotes or suppresses the reproduction of
XX hepatitis C virus, involving culturing the replicon reproducing cell in
XX the presence of a test substance, and detecting the reproduction of
XX replicon RNA in the culture. Virucide. The replicon RNA is useful for
XX producing a replicon reproduction cell and for increasing the
XX reproduction efficiency of replicon RNA of hepatitis C virus of genotype
XX 2a. The cell and the replicon RNA are useful for producing a therapeutic
XX agent or a diagnostic agent for hepatitis C virus infection, for
XX producing a vaccine against hepatitis C virus infection and for screening
XX a substance that promotes or suppresses the reproduction of hepatitis C
XX virus. This sequence represents a hepatitis C virus (HCV) protein used in
XX the scope of the invention.
XX
SQ Sequence 3033 AA;

Query Match 5.9%; Score 7; DB 9; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 959
ADV04738
ID ADV04738 standard; protein; 3033 AA.
XX
AC ADV04738;
XX
DT 24-FEB-2005 (first entry)
XX
DE Hepatitis C virus (HCV) protein #1.
XX
KW Virucide; hepatitis C virus infection; replicon.
XX
OS Hepatitis C virus.
XX
FN WO2004104198-A1.
XX
PD 02-DEC-2004.
XX
PF 25-NOV-2003; 2003WO-JP015038.

CC response against variants of the peptide epitope. This sequence
CC represents an HCV polymerase protein used in the scope of the invention.
XX Sequence 3033 AA;

Query Match 5.9%; Score 7; DB 9; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
Db 1700 PDKEVLY 1706

RESULT 961

ADX40778
ID ADX40778 standard; protein; 3037 AA.

XX AC ADX40778;

XX DT 21-APR-2005 (first entry)

XX DE HCV polymerase protein #1.

XX KW Immune stimulation; polymerase; enzyme.

XX OS Hepatitis C virus.

XX PN WO2005012502-A2.

XX PD 10-FEB-2005.

XX PF 29-MAR-2004; 2004WO-US009510.

XX PR 28-MAR-2003; 2003US-0458026P.

XX PA (EPIM-) EPIMUNE INC.

XX PI Baker DM, Livingston BD, Chesnut RW, Sette A, Newman MJ;
XX WPI; 2005-132661/14.

XX PT Identifying a candidate peptide epitope, which induces a HLA class I CTL
XX response comprises identifying variants of a peptide epitope 8-11 amino
XX acids in length comprising primary anchor residues of the same HLA class
XX I binding motif.

XX PS Disclosure; Page 387-440; 458pp; English.

XX CC The invention relates to a method of identifying a candidate peptide
XX epitope which induces an HLA class I CTL response against variants of the
XX peptide epitope, comprising identifying, from a particular antigen of an
XX infectious agent, variants of a peptide epitope comprising primary anchor
XX residues of the same HLA class I binding motif. The method is useful for
XX identifying a candidate peptide epitope, which induces an HLA class I CTL
XX response against variants of the peptide epitope. This sequence
XX represents an HCV polymerase protein used in the scope of the invention.

XX SQ Sequence 3037 AA;

Query Match 5.9%; Score 7; DB 9; Length 3037;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
Db 1700 PDKEVLY 1706

RESULT 962

AAV53675
ID AAV53675 standard; protein; 3262 AA.

XX

AC AAV53675;

XX DT 22-FEB-2000 (first entry)

XX DE Mechanical stress induced protein 274 amino acid sequence.

XX KW Mechanical stress; gene therapy; protein 274; osteoporosis; bone density;
XX bone development.

XX OS Rattus sp.

XX FH Key Location/Qualifiers

XX FT Misc-difference 115 /note= "unknown amino acid encoded by SCA"

XX FT Misc-difference 1608 /note= "encoded by CGT"

XX FT Misc-difference 1632 /note= "encoded by TTC"

XX FT Misc-difference 1632 /note= "encoded by TTC"

XX PN WO9960164-A1.

XX PD 25-NOV-1999.

XX PF 14-MAY-1999; 99WO-US011066.

XX PR 15-MAY-1998; 98US-0085673P.

XX PA (QUAR-) QUARK BIOTECH INC.

XX PI Einat P, Mor O, Skalliter R, Feinstein E, Faerman A;

XX DR WPI; 2000-053304/04.

XX DR N-PSDB; AAZ36325.

XX PT Identification of stress induced genes for determining risk and
XX preventing, treating or controlling osteoporosis.

XX PS Claim 32; Fig 13; 308pp; English.

XX CC The present sequence represents protein 274, which was identified using
XX the method of the invention after subjecting rat osteoblasts to
XX mechanical stress. Expression of the 274 gene was found to be upregulated
XX by about 3-fold in cells subjected to mechanical strain. The
XX specification describes a method for the identification of genes
XX responsive to a specific mechanical stress. The method comprises applying
XX the mechanical stress to an organism (tissue or cells comprising bone
XX cells), isolating the specific cellular fractions and extracting mRNA
XX from them, and differentially analysing the mRNA in comparison with
XX control samples. The method is used to identify genes whose expression is
XX responsive to a specific stress. The identified genes are employed in
XX determining risk associated with a physiological or disease state. The
XX risk determination methods are used for testing a medication for gene
XX therapy. These medications, or genes identified by the method of the
XX invention, are used for treating, preventing or controlling a
XX physiological or disease state (especially osteoporosis or bone density
XX or other factors causing or contributing to osteoporosis or its symptoms
XX or other conditions involved in mechanical stress or its lack. The
XX methods can also be used for advancing research or studies in bone
XX development

XX SQ Sequence 3262 AA;

XX Query Match 5.9%; Score 7; DB 3; Length 3262;

XX Best Local Similarity 100.0%; Pred. No. 3e+03;

XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 20 VLAALAA 26
|||||

XX Db 1479 VLAALAA 1485

XX RESULT 963

XX AAV53676

```

ID XX AAY53676 standard; protein; 3264 AA.
AC XX AAY53676;
XX DT
XX DT 22-FEB-2000 (first entry)
XX DE Protein 274 sequence used for clustral X alignment.
XX KW Mechanical stress; gene therapy; protein 274; osteoporosis; bone density;
XX KW bone development.
XX OS Rattus sp.
XX FH
XX FT Key Location/Qualifiers
XX FT Misc-difference 1862
XX FT /note= "unspecified amino acid"
XX PN
XX XX MO9960164-A1.
XX PD
XX XX 25-NOV-1999.
XX XX 14-MAY-1999; 99WO-US011066.
XX XX 15-MAY-1998; 98US-0085673P.
XX XX (QUAR-) QUARK BIOTECH INC.
XX PA
XX PI Einat P, Mor O, Skalter R, Feinstein E, Faerman A;
XX XX WPI; 2000-053304/04.
XX DR
XX XX Identification of stress induced genes for determining risk and
XX FT preventing, treating or controlling osteoporosis.
XX P8
XX XX Claim 32; Fig 14A-T; 308pp; English.
XX CC The present sequence is a fragment of protein 274, used for clustral X
XX CC alignment. Protein 274 was identified using the method of the invention
XX CC after subjecting rat osteoblasts to mechanical stress. Expression of the
XX CC 608 gene was found to be upregulated by about 3-fold in cells subjected
XX CC to mechanical strain. The specification describes a method for the
XX CC identification of genes responsive to a specific mechanical stress. The
XX CC method comprises applying the mechanical stress to an organism (tissue or
XX CC cells comprising bone cells), isolating the specific cellular fractions
XX CC and extracting mRNA from them, and differentially analysing the mRNA in
XX CC comparison with control samples. The method is used to identify genes
XX CC whose expression is responsive to a specific stress. The identified genes
XX CC are employed in determining risk associated with a physiological or
XX CC disease state. The risk determination methods are used for testing a
XX CC medicament for gene therapy. These medicaments, or genes identified by
XX CC the method of the invention, are used for treating, preventing or
XX CC controlling a physiological or disease state (especially osteoporosis or
XX CC bone density or other factors causing or contributing to osteoporosis or
XX CC its symptoms or other conditions involved in mechanical stress or its
XX CC bone development
XX CC bone development
XX SQ Sequence 3264 AA;
XX Query Match 5.9%; Score 7; DB 3; Length 3264;
XX Best Local Similarity 100.0%; Pred. No. 3e+03;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALAA 26
DB 787 VLAALAA 793
|||||
RESULT 964
AAE14793
ID AAE14793 standard; protein; 5183 AA.
XX XX
XX AC AAE14793;

XX DT
XX DE Human microtubule-associated factor (MTAF) 600.
XX KW Microtubule-associated factor 600; MTAF600; retinoblastoma; RB;
XX KW tumour suppressor; E2F; transcriptional activator; cell cycle;
XX KW microtubule; hyperproliferative condition; retinoblastoma; osteosarcoma;
XX KW cancer; lung; breast; bladder; human; gene therapy.
XX OS Homo sapiens.
XX FH
XX FT Key Location/Qualifiers
XX FT Domain 1650..1730
XX FT /label= Zinc_finger_domain
XX FT Binding-site 3910..4851
XX FT /note= RB binding site
XX FT /note= "MTAF600 binds to large pocket domain residues 379
XX FT -928) retinoblastoma (RB) polypeptide"
XX FT Binding-site 4076..4122
XX FT /label= Calmodulin_(CAM)_binding_domain
XX FT Binding-site 4091..4113
XX FT /label= CAM_binding_site
XX FT Binding-site 4293..4534
XX FT /label= RB_binding_site
XX FT /note= "MTAF600 binds to large pocket domain residues 379
XX FT -928) retinoblastoma (RB) polypeptide"
XX PN
XX XX WO200277019-A1.
XX XX 03-OCT-2002.
XX XX 25-MAR-2002; 2002WO-US009382.
XX XX 23-MAR-2001; 2001US-0278244P.
XX XX 23-MAR-2001; 2001US-0278245P.
XX XX (DAND ) DANA FARBER CANCER INST INC.
XX XX Nakatani Y;
XX WPI; 2003-018883/01.
XX CC New microtubule associated factor 600 protein, useful for preparing a
XX CC composition for treating cancer e.g., retinoblastoma, osteosarcoma, or
XX CC lung, breast or bladder cancer.
XX CC Claim 1; Fig 1b; 75pp; English.
XX CC The invention relates to microtubule-associated factor 600 (MTAF600), a
XX CC 600kDa subunit of retinoblastoma (RB) polypeptide complex. RB is a tumour
XX CC suppressor and its major role is repression of E2F family of DNA-binding
XX CC transcriptional activators, which regulate cell cycle through various
XX CC genes required for S-phase entry. MTAF600 interacts directly with RB and
XX CC microtubules and plays a role in active repression of E2F-responsive
XX CC genes, cell cycle arrest and genomic stability. The pharmaceutical
XX CC composition comprising the polypeptide of the invention is useful for
XX CC treating hyperproliferative condition, particularly cancer such as
XX CC retinoblastoma, osteosarcoma, lung, breast or bladder cancer. The present
XX CC sequence is human MTAF600 polypeptide
XX SQ Sequence 5183 AA;
XX Query Match 5.9%; Score 7; DB 6; Length 5183;
XX Best Local Similarity 100.0%; Pred. No. 4.5e+03;
XX Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALAA 26
DB 3332 VLAALAA 3338
|||||
RESULT 965
AAE14793
```

AD044006
ID AD044006 standard; protein; 5183 AA.
XX AC
XX AD044006;
XX DT 15-JUL-2004 (first entry)
XX DE Amino acid sequence of human Pushover.
XX KW protein complex; neurological disease; stroke; neurodegeneration;
XX KW Wallerian degeneration; Alzheimer's disease; neurological disorder;
XX KW epilepsy; inflammatory condition; ulcerative colitis; Crohn's disease;
XX KW atherosclerosis; ID-MYO-inositol triphosphate 3 kinase A; ASK1; ASK2;
XX KW ASK3; CAMKII beta; CAMKII delta; CAMKII gamma; casein kinase II alpha;
XX KW Cdc37; CHK2; CTCL tumour antigen SE20-4; EP-1 alpha 1; ENAP;
XX KW FLJ14653 NT2RP2002252; FLJ30839 FBRA2002429; HERC2;
XX KW inositol polyphosphate-5-phosphatase; inositol-1; 4;
XX KW 5-triphosphate 5-phosphatase type 1; IRAK1; IRAK4; KIAA1441; MSTP030;
XX KW Nek9; PAR3; Pellino 1; Pellino 3; podocalyxin-like protein 1 precursor;
XX KW Pushover; S-adenosylhomocysteine;
XX KW secretory carrier-associated membrane protein 2; surfactant protein 2;
XX KW ubiquitin carboxyl terminal hydrolase 11;
XX KW upstream regulatory element binding protein 1; Vartul;
XX KW Werner's syndrome helicase interacting protein; WHIP;
XX KW X-ray repair cross complementing protein 4.
XX OS Homo sapiens.
XX PN WO2004031242-A2.
XX PD 15-APR-2004.
XX PP 11-SEP-2003; 2003WO-EP010110.
XX PR 12-SEP-2002; 2002EP-00020495.
XX PR 12-SEP-2002; 2002EP-00020496.
XX PR 12-SEP-2002; 2002EP-00020497.
XX PA (CELL-) CELLZOME AG.
XX PI Bouwmeester T, Drewes G, Jackson D, Helftenbein G, Schirle M;
XX PI Kuester B, Hopf C;
XX WPI; 2004-316467/29.
XX New complex comprising at least one first protein, and at least one
XX second protein, useful for treating stroke, Alzheimer's disease.
XX neurological disorders such as epilepsy, and inflammatory conditions such
XX as ulcerative colitis.
XX Example; Page 226-242; 287pp; English.
XX The specification describes protein complexes involved in cellular
XX processes which have been shown to be critical for the development of
XX various forms of neurological diseases. Three protein complexes were
XX identified: ASK2 protein complex, Pellino-1 protein complex and Pellino-3
XX protein complex. The protein complex are useful for treating diseases and
XX disorders, e.g. stroke, neurodegeneration such as Wallerian degeneration,
XX Alzheimer's disease, neurological disorders such as epilepsy, and
XX inflammatory conditions such as ulcerative colitis, Crohn's disease or
XX atherosclerosis. Proteins identified as being part of the protein
XX complexes of the invention are ID-MYO-inositol triphosphate 3 kinase A,
XX ASK1, ASK2, ASK3, CAMKII beta, CAMKII delta, CAMKII gamma, casein kinase
XX II alpha, Cdc37, CHK2, CTCL tumour antigen SE20-4, EP-1 alpha 1, ENAP,
XX FLJ14653 NT2RP2002252, FLJ30839 FBRA2002429, HERC2, two hypothetical
XX proteins of 35.5 kDa and 49.3 kDa, inositol polyphosphate-5-phosphatase,
XX inositol-1,4,5-triphosphate 5-phosphatase type 1, IRAK1, IRAK4, KIAA1441,
XX MSTP030, Nek9, PAR3, Pellino 1, Pellino 3, podocalyxin-like protein 1
XX precursor, Pushover, a putative S-adenosylhomocysteine, secretory
XX carrier-associated membrane protein 2, surfactant protein 2, ubiquitin
XX carboxyl terminal hydrolase 11, upstream regulatory element binding
XX protein 1, Vartul, Werner's syndrome helicase interacting protein (WHIP),
XX X-ray repair cross complementing protein 4 (isoform 1). The present

CC sequence represents Pushover.
XX SQ Sequence 5183 AA;
Query Match 5.9%; Score 7; DB 8; Length 5183;
Best Local Similarity 100.0%; Pred. No. 4.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAAALAA 26
Db 3332 VLAAALAA 3338
RESULT 966
ADX06154
ID ADX06154 standard; protein; 5183 AA.
XX AC ADX06154;
XX DT 21-APR-2005 (first entry)
XX DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 719.
XX KW cytostatic; cyclin-dependent kinase; cdk; biomarker.
XX OS Homo sapiens.
XX PN WO2005012875-A2.
XX PD 10-FEB-2005.
XX PF 29-JUL-2004; 2004WO-US024424.
XX PR 29-JUL-2003; 2003US-0490890P.
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
XX WPI; 2005-163068/17.
XX N-PSDB; ADX06153.
XX Biomarkers useful for predicting or determining the response of a mammal
XX to a cancer treatment comprising administration of a modulator of cyclin-
XX dependent kinase activity.
XX Claim 5; SEQ ID NO 719; 141pp; English.
XX This invention describes a novel method of predicting or determining
XX whether a mammal will respond or is responding to an anti-cancer agent
XX that modulates cyclin-dependent kinase (cdk) activity. The method agent
XX comprises measuring the level of one or more biomarkers selected from
XX 2774 biomarkers given in the specification (nucleotide sequence SEQ ID
XX NO:1246 (Genbank EST W28729) is especially preferred). The method of the
XX invention is utilized in a kit for determining or predicting whether
XX patient would be susceptible or resistant to treatment by an agent
XX modulating cdk activity. The invention also describes a method for
XX utilizing individualized genetic profiles for treating diseases and
XX disorders based on patient's response and molecular level, specialized
XX microarrays comprising the biomarkers described, antibodies directed
XX against the biomarkers and a cell culture model to identify biomarkers.
XX The cdk modulator is preferably N-5-[(5-(1,1-dimethylethyl)-2-
XX oxazolyl)methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-L-
XX tartaric acid salt. Note: The sequence data for this patent did not form
XX part of the printed specification, but was obtained in electronic format
XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This
XX sequence represents a biomarker used in the method of the invention.
XX SQ Sequence 5183 AA;
Query Match 5.9%; Score 7; DB 9; Length 5183;
Best Local Similarity 100.0%; Pred. No. 4.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      20 VLALAA 26
Db      3332 VLALAA 3338
      |||||
RESULT 967
AAE36120
ID   AAE36120 standard; protein; 5644 AA.
AC   AAE36120;
XX
XX      26-JUN-2003 (first entry)
XX
XX      Streptomyces nodosus amphJ gene encoded protein.
XX
XX      Polyene; antibiotic; amphotericin; amph; polyketide synthase; enzyme.
XX
XX      Streptomyces nodosus.
XX
XX      WO200297082-A2.
XX
XX      05-DEC-2002.
XX
XX      27-MAY-2002; 2002WO-IE0000071.
XX
XX      31-MAY-2001; 2001IE-00000527.
XX
XX      (UYDU-) UNIV COLLEGE DUBLIN.
XX
XX      Caffrey JP;
XX
XX      WPI; 2003-201271/19.
XX      N-PSDB; AAD54645.
XX
XX      Novel cytochrome P450 enzyme and nucleotides encoding the enzyme, useful
XX      for preparing amphotericin derivative or analog antibiotic agent with
XX      altered properties, in biosynthesis of polyketide other than
XX      amphotericin.
XX
XX      Claim 6; Page 162-187; 276pp; English.
XX
XX      The invention relates to the gene cluster encoding the polypeptides
XX      responsible for the biosynthesis of the polyene antibiotic amphotericin
XX      (amph) of Streptomyces nodosus. Polynucleotides of the invention are
XX      useful for preparing amphotericin derivatives or analogue antibiotic
XX      agents with altered properties and in the biosynthesis of polyketides
XX      other than amphotericin. amphDII, amphDII or amphDI mutants are useful
XX      for producing amphotericin derivatives glycosylated with alternative
XX      sugars; amphDII or amphDII gene sequences are useful in engineered
XX      biosynthesis of perosaminyl-amphoteronolide B; amphDII or amphDII and
XX      amphN gene sequences are useful in the engineered biosynthesis of
XX      perosaminyl-16-desacboxyl-16-methyl amphoteronolide B; amphDIII, amphDII
XX      and amphDI gene sequences are useful for preparing polypeptides capable
XX      of addition of mycosamine to a polyketide other than amphoteronolide A or
XX      B or for preparing polypeptides for in vitro synthesis of GDP-mycosamine.
XX      The present sequence is polyketide synthase multienzyme housing extension
XX      modules 15, 16 and 17 encoded by S. nodosus amphJ gene
XX
XX      Sequence 5644 AA;
XX
XX      Query Match      5.9%; Score 7; DB 6; Length 5644;
XX      Best Local Similarity 100.0%; Pred. No. 4.8e+03;
XX      Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      19 GVLAALA 25
Db      851 GVLAALA 857
      |||||
RESULT 968
ABU11385
ID   ABU11385 standard; protein; 7349 AA.

```

```

XX      ABU11385;
XX
XX      11-FEB-2003 (first entry)
XX
XX      Protein encoded by S. atroolivaceus leinamycin gene cluster ORF lnmJ.
XX
XX      Leinamycin biosynthesis gene cluster; lnm; open reading frame; ORF;
XX      anti-tumour antibiotic; broad spectrum antimicrobial activity;
XX      Gram-positive; Gram-negative bacteria; chemical modification; metabolite;
XX      apo-carrier protein; holo-carrier protein; tumour; polyketide;
XX      hybrid polypeptide/polyketide metabolite; lnm production; cytostatic.
XX
XX      Streptomyces atroolivaceus.
XX
XX      WO200277179-A2.
XX
XX      03-OCT-2002.
XX
XX      22-MAR-2002; 2002WO-US008937.
XX
XX      26-MAR-2001; 2001US-0278935P.
XX      (REGC ) UNIV CALIFORNIA.
XX      (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX      Shen B, Cheng Y, Tang G;
XX      WPI; 2003-018907/01.
XX      N-PSDB; ABX34289.
XX
XX      Novel gene cluster responsible for synthesis of leinamycin in
XX      Streptomyces atroolivaceus useful for making various peptide and/or
XX      polyketide, and/or hybrid polypeptide/polyketide metabolites.
XX
XX      Claim 13; Page 145-149; 185pp; English.
XX
XX      The present invention relates to the isolation of the Streptomyces
XX      atroolivaceus leinamycin (lnm) biosynthesis gene cluster containing 71
XX      open reading frames (ORFs) (ORFs -35 through -1, ORFs lnmA through lnmZ,
XX      and ORFs +1 through +9). Leinamycin is a novel anti-tumour antibiotic
XX      produced by several Streptomyces species. It exhibits broad spectrum
XX      antimicrobial activity against Gram-positive and Gram-negative bacteria,
XX      but not against fungi. The polypeptides encoded by the lnm biosynthesis
XX      gene cluster ORFs are useful for chemically modifying a molecule in a
XX      host cell. The host cell is a bacterium or eukaryotic cell, including a
XX      mammalian, yeast, plant, fungal, or insect cell. The molecule is an
XX      endogenous metabolite produced by the host cell or exogenously supplied
XX      metabolite, or an amino acid, and the polypeptide is a peptide synthetase
XX      or amino transferase. The polypeptides encoded by the lnm gene cluster
XX      are useful for converting an apo-carrier protein to a holo-carrier
XX      protein. lnm shows potent antitumour activity in tumour models in vivo.
XX      The lnm gene cluster modules and/or catalytic domains are useful for
XX      making various peptide and/or polyketide, and/or hybrid
XX      polypeptide/polyketide metabolites. The proteins encoded by the ORFs are
XX      useful alone, or in combination with other active domains to modify
XX      various target substrates. The lnm gene cluster is useful to upregulate
XX      endogenous lnm production to permit lnm production in cells and/or to
XX      make various modified lnm, lnm, its analogue, or other polyketide,
XX      peptide or hybrid polyketide/peptide metabolites are useful as
XX      therapeutic agents, to treat a number of disorders, depending upon the
XX      type of metabolites. ABU11341-ABU11411 represent the proteins encoded by
XX      ORFs of the S. atroolivaceus leinamycin biosynthesis gene cluster
XX
XX      Sequence 7349 AA;
XX
XX      Query Match      5.9%; Score 7; DB 6; Length 7349;
XX      Best Local Similarity 100.0%; Pred. No. 6e+03;
XX      Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY      20 VLALAA 26
Db      3463 VLALAA 3469
      |||||

```

RESULT 969
 AAY28715
 ID AAY28715 standard; peptide; 6 AA.
 XX AC AAY28715;
 XX AC AAY28715;
 XX DT 27-AUG-2003 (revised)
 XX DT 01-NOV-1999 (first entry)
 XX Human hepatitis C virus NS3 serine protease peptide inhibitor-2.
 XX Human hepatitis C virus NS3 serine protease peptide inhibitor-2; HCV;
 KW NS3 protein; HCV polypeptide non-structural portion; NS3 protease; IC50;
 KW NS3/NS4A junction site; NS4A/NS4B junction site; NS4B/NS5A junction site;
 KW NS5A/NS5B junction site; HCV NS3 protease binding assay; drug;
 XX HCV NS3 protease inhibition assay; non-A non-B hepatitis.
 OS Synthetic.
 OS Hepatitis C virus.
 XX Key Location/Qualifiers
 FT Misc-difference 6 /note= "D-form residue"
 FT
 XX WO9938888-A2.
 XX 05-AUG-1999.
 XX 02-FEB-1999; 99WO-IT0000022.
 XX 02-FEB-1999; 98IT-RM0000061.
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX Pessi A, Steinkuehler C, De Francesco R;
 XX WPI; 1999-479167/40.
 XX Hepatitis C virus NS3 serine protease peptide inhibitors, useful for
 treatment of non-A non-B hepatitis.
 XX Claim 19; Page 10; 82pp; English.
 XX The present sequence is a peptide inhibitor of serine protease activity
 associated to the NS3 protein of human hepatitis C virus (HCV). The
 peptide is a result of modification of one of the peptide sequences
 obtained by proteolytic action of NS3 protease on HCV polypeptide
 comprising junction sites selected from NS3/NS4A, NS4A/NS4B, NS4B/NS5A
 and NS5A/NS5B. The peptide is capable of inhibiting 50% of the NS3
 protease activity at a concentration of 4.0 micro molar (IC50). The
 inhibitor is used for derivation of binding or inhibition assays of the
 HCV NS3 protease and preparation of drugs for non-A non-B hepatitis.
 XX (Updated on 27-AUG-2003 to correct OS field.)
 XX Sequence 6 AA;
 Query Match 5.1%; Score 6; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMEEC 65
 DB 1 DEMEEC 6
 RESULT 970
 AAY28714
 ID AAY28714 standard; peptide; 6 AA.
 XX AC AAY28714;
 XX AC AAY28714;

DT 27-AUG-2003 (revised)
 DT 01-NOV-1999 (first entry)
 XX Human hepatitis C virus NS3 serine protease peptide inhibitor-1.
 DE Human hepatitis C virus NS3 serine protease peptide inhibitor-1; HCV;
 KW NS3 protein; HCV polypeptide non-structural portion; NS3 protease; IC50;
 KW NS4A/NS4B junction site; HCV NS3 protease binding assay; drug;
 KW HCV NS3 protease inhibition assay; non-A non-B hepatitis.
 XX Hepatitis C virus.
 OS WO9938888-A2.
 XX 05-AUG-1999.
 XX 02-FEB-1999; 99WO-IT0000022.
 XX 02-FEB-1999; 98IT-RM0000061.
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
 XX Pessi A, Steinkuehler C, De Francesco R;
 XX WPI; 1999-479167/40.
 XX Hepatitis C virus NS3 serine protease peptide inhibitors, useful for
 treatment of non-A non-B hepatitis.
 XX Claim 19; Page 10; 82pp; English.
 XX The present sequence is a peptide inhibitor of serine protease activity
 associated to the NS3 protein of human hepatitis C virus (HCV). The
 peptide corresponds to a part of the sequence present at the junction of
 NS4A and NS4B proteins of the HCV polypeptide and is obtained by the
 proteolytic action of NS3 protease on the NS4A/NS4B junction site of HCV
 polypeptide. The peptide is capable of inhibiting 50% of the NS3 protease
 activity at a concentration of 1.0 micro molar (IC50). The inhibitor is
 used for derivation of binding or inhibition assays of the HCV NS3
 protease and preparation of drugs for non-A non-B hepatitis. (Updated on
 27-AUG-2003 to correct OS field.)
 XX Sequence 6 AA;
 Query Match 5.1%; Score 6; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 DEMEEC 65
 DB 1 DEMEEC 6
 RESULT 971
 AAY28740
 ID AAY28740 standard; peptide; 6 AA.
 XX AC AAY28740;
 XX 27-AUG-2003 (revised)
 XX 01-NOV-1999 (first entry)
 XX Human hepatitis C virus NS3 serine protease peptide inhibitor-27.
 DE Human hepatitis C virus NS3 serine protease peptide inhibitor-27; HCV;
 KW NS3 protein; HCV polypeptide non-structural portion; NS3 protease; IC50;
 KW NS3/NS4A junction site; NS4A/NS4B junction site; NS4B/NS5A junction site;
 KW NS5A/NS5B junction site; HCV NS3 protease binding assay; drug;
 XX HCV NS3 protease inhibition assay; non-A non-B hepatitis.
 OS Synthetic.
 OS Hepatitis C virus.
 XX

PT Peptide compounds useful for treating or preventing hepatitis C
PT infections bind to hepatitis C virus NS3 protease.
XX
XX
PS Disclosure; Page 14; 46pp; English.
XX
CC The present sequence is one of a large number of peptidic inhibitors of
CC hepatitis C virus NS3 serine protease. HCV infection can result in
CC chronic hepatitis and cirrhosis of the liver, and may lead to
CC hepatocellular carcinoma. Currently no vaccine nor established therapy
CC exists. The present peptide is a reversible, non-covalent inhibitor that
CC is based on the P and P' regions of the natural substrate of NS3. The P'
CC part of the inhibitor is optimised to achieve maximum binding energy
CC through interaction with the S' region of the enzyme. The peptides were
CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
CC The ability of the compounds to inhibit NS3 protease was evaluated using
CC a complex comprising the NS3 protease domain and a modified form of the
CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
CC of the disclosed peptides to inhibit cleavage of the substrate by the
CC complex was measured and some of the peptides were found to be potent
CC inhibitors of HCV protease. The compounds can be used at micromolar and
CC nanomolar levels to treat or prevent hepatitis C or a related condition
XX
SQ Sequence 6 AA;
Query Match 5.1%; Score 6; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMERC 65
DB 1 DEMERC 6
RESULT 974
AA97437
ID AA97437 standard; peptide; 6 AA.
XX
AC AA97437;
XX
DT 11-SEP-2000 (first entry)
XX
DE Hepatitis C virus NS3 protease peptidic inhibitor #2.
XX
KW Hepatitis C virus; HCV; serine protease; virucide; infection;
KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
KW hepatocellular carcinoma.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT Misc-difference 6 /note= "D-form residue"
FT
XX WO200031129-A1.
XX
XX 02-JUN-2000.
XX
XX 24-NOV-1999; 99WO-EP009207.
XX
XX 26-NOV-1998; 98GB-00025946.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX Pessi A, Ingallinella P, Bianchi E;
XX WPI; 2000-411933/35.
XX
XX Peptide compounds useful for treating or preventing hepatitis C
XX infections bind to hepatitis C virus NS3 protease.

PS Disclosure; Page 14; 46pp; English.
XX
XX The present sequence is one of a large number of peptidic inhibitors of
XX hepatitis C virus NS3 serine protease. HCV infection can result in
XX chronic hepatitis and cirrhosis of the liver, and may lead to
XX hepatocellular carcinoma. Currently no vaccine nor established therapy
XX exists. The present peptide is a reversible, non-covalent inhibitor that
XX is based on the P and P' regions of the natural substrate of NS3. The P'
XX part of the inhibitor is optimised to achieve maximum binding energy
XX through interaction with the S' region of the enzyme. The peptides were
XX synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
XX The ability of the compounds to inhibit NS3 protease was evaluated using
XX a complex comprising the NS3 protease domain and a modified form of the
XX NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
XX the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
XX of the disclosed peptides to inhibit cleavage of the substrate by the
XX complex was measured and some of the peptides were found to be potent
XX inhibitors of HCV protease. The compounds can be used at micromolar and
XX nanomolar levels to treat or prevent hepatitis C or a related condition
XX
SQ Sequence 6 AA;
Query Match 5.1%; Score 6; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMERC 65
DB 1 DEMERC 6
RESULT 975
AA97436
ID AA97436 standard; peptide; 6 AA.
XX
AC AA97436;
XX
DT 11-SEP-2000 (first entry)
XX
DE Hepatitis C virus NS3 protease peptidic inhibitor #1.
XX
KW Hepatitis C virus; HCV; serine protease; virucide; infection;
KW virus protease inhibitor; chronic hepatitis; liver cirrhosis;
KW hepatocellular carcinoma.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT
XX WO200031129-A1.
XX
XX 02-JUN-2000.
XX
XX 24-NOV-1999; 99WO-EP009207.
XX
XX 26-NOV-1998; 98GB-00025946.
XX
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX Pessi A, Ingallinella P, Bianchi E;
XX WPI; 2000-411933/35.
XX
XX Peptide compounds useful for treating or preventing hepatitis C
XX infections bind to hepatitis C virus NS3 protease.
XX
XX Disclosure; Page 14; 46pp; English.
XX
XX The present sequence is one of a large number of peptidic inhibitors of
XX hepatitis C virus NS3 serine protease. HCV infection can result in
XX chronic hepatitis and cirrhosis of the liver, and may lead to

CC hepatocellular carcinoma. Currently no vaccine nor established therapy
CC exists. The present peptide is a reversible, non-covalent inhibitor that
CC is based on the P and P' regions of the natural substrate of NS3. The P'
CC part of the inhibitor is optimised to achieve maximum binding energy
CC through interaction with the S' region of the enzyme. The peptides were
CC synthesised on solid phase by the continuous-flow Fmoc-polyamide method.
CC The ability of the compounds to inhibit NS3 protease was evaluated using
CC a complex comprising the NS3 protease domain and a modified form of the
CC NS4A peptide, Pep 4AK. A substrate (peptide 4AB) based on the sequence of
CC the NS4A/NS4B cleavage site of the HCV polyprotein was used. The ability
CC of the disclosed peptides to inhibit cleavage of the substrate by the
CC complex was measured and some of the peptides were found to be potent
CC inhibitors of HCV protease. The compounds can be used at micromolar and
CC nanomolar levels to treat or prevent hepatitis C or a related condition
XX
SQ Sequence 6 AA;

Query Match 5.1%; Score 6; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEEC 65
Db 1 DEMEEC 6
|||||

RESULT 976
AAO23377
ID AAO23377 standard; peptide; 6 AA.

XX
AC AAO23377;

XX
DT 06-NOV-2003 (first entry)

XX
DE HCV NS4A peptide 12 (SeqID 28) used to identify binding to NS3 protease.

XX
KW Hepatitis C virus; HCV; serine protease inhibitor; sequellae; cirrhosis;
KW chronic active hepatitis; hepatocellular carcinoma; liver fibrosis; NS3;
KW necrosis; inflammation; bile duct change; cytostatic; antiinflammatory;
KW bovine viral diarrhoea; hepatotropic; antiviral; virucide; dengue fever;
KW NS4A.

XX
OS Hepatitis C virus.

XX
FN WO2003051910-A2.

XX
PD 26-JUN-2003.

XX
PF 13-DEC-2002; 2002WO-CA001929.

XX
PR 14-DEC-2001; 2001US-0340574P.

XX
PA (JOYC/) JOYCE M.

XX
PA (WILL/) WILLIAMS M.

XX
PA (HIND/) HINDSGAUL O.

XX
PA (TYRR/) TYRREL D L.

XX
PI Joyce M, Williams M, Hindsgaul O, Tyrrel DL;

XX
DR WPI; 2003-607859/57.

XX
PT New peptides useful, e.g. in the treatment of or reduction of viral load
PT of hepatitis C virus and associated conditions, e.g. liver fibrosis,

XX
PS Example 3; Fig 7; 30pp; English.

XX
CC This invention relates to novel hepatitis C virus (HCV) protease
CC inhibitors. Specifically, these inhibitors are small, hydrophobic
CC peptides that work by affecting the activity of the HCV serine protease
CC NS3, or preventing NS3 activation by inhibition of its co-factor NS4A.
CC Chronic infection with HCV can lead to serious sequellae including
CC Chronic active hepatitis, cirrhosis and hepatocellular carcinoma, as well

CC as HCV associated conditions including liver fibrosis, necrosis,
CC inflammation or bile duct changes. The present invention describes these
CC peptide inhibitors as virucides, and as such they can be used to inhibit
CC HCV replication and reduce the viral load. They also have hepatotropic
CC and antiinflammatory activity and can be described as cytostatic.
CC Furthermore, the antiviral peptides derived from the relevant conserved
CC NS3 or NS4A domains can be used to treat other viruses including the
CC dengue fever virus and the bovine viral diarrhoea virus. This peptide
CC sequence, peptide 12 (SeqID 28), is part of the C-terminal deletion
CC library that was used to identify the minimal domain of NS4A that is
CC required for binding to the NS3 protease, and hence identify inhibitor
XX
XX peptides of the invention
SQ Sequence 6 AA;

Query Match 5.1%; Score 6; DB 6; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIV 37
Db 1 GCVVIV 6
|||||

RESULT 977

ADU39858
ID ADU39858 standard; peptide; 6 AA.

XX
AC ADU39858;

DT 27-JAN-2005 (first entry)

XX
DE Peptide inhibitor (HCV Inh 1).

XX
KW Antiinflammatory; hepatotropic; virucide;

XX
KW hepatitis C serine protease inhibitor;

XX
KW hepatitis C virus replication inhibitor; NS3 protease inhibitor;

XX
KW hepatitis C virus; helicase inhibitor; polymerase inhibitor;

XX
KW metalloprotease inhibitor; internal ribosome entry site inhibitor; IRES.

XX
OS Synthetic.

XX
FH Key Location/Qualifiers

XX
FT Modified-site 1 /note= "Ac-Asp-"

XX
FN WO2004093798-A2.

XX
PD 04-NOV-2004.

XX
PF 16-APR-2004; 2004WO-US011841.

XX
PR 18-APR-2003; 2003US-00418759.

XX
PA (ENAN-) ENANTA PHARM INC.

XX
PI Nakajima S, Sun Y, Tang D, Xu G, Porter B, Or YS, Wang Z;

XX
PI Miao Z;

XX
XX WPI; 2004-795391/78.

XX
DR Example 81; Page 110; 131pp; English.

XX
CC The invention relates to novel quinoxaliny macrocyclic derivatives (I)
CC and (II). Further disclosed is a method to inhibit the replication of
CC hepatitis C virus, comprising supplying a hepatitis C viral NS3 protease
CC inhibitory amount of a compound of the invention. The method further
CC comprises concurrent administration of an additional antihepatitis C
CC virus agent (alpha-interferon, beta-interferon, ribavirin or adamantane).

CC The anti-hepatitis agent is an inhibitor of hepatitis C virus helicase,
 CC polymerase, metalloprotease or internal ribosome entry site (IRES).
 CC Compounds of the invention are useful to treat a hepatitis C viral
 CC infection and to inhibit the replication of hepatitis C virus. Compounds
 CC of the invention have increased biological penetration into a given
 CC biological system e.g. blood or lymphatic system, increased oral
 CC availability, increased solubility to allow administration by injection,
 CC alter metabolism and alter rate of excretion. The current sequence
 CC represents a peptide inhibitor (HCV Inh 1) used in an example from the
 CC invention as a reference compound in a NS3/NS4a protease enzyme assay.
 SQ Sequence 6 AA;

Query Match 5.1%; Score 6; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
 Db 1 DEMEEC 6
 |||||

RESULT 978

ADV78811
 ID ADV78811 standard; peptide; 6 AA.

XX AC ADV78811;

XX XX 10-MAR-2005 (first entry)

XX DE HCV inhibitory peptide Inh1.

XX KW Hepatitis C virus infection; gastrointestinal disease; infection;
 KW drug screening; Virucide; Antiinflammatory; Hepatotropic;
 KW serine protease inhibitor; pharmaceutical.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 XX FT Modified-site 1 /note= "Acetylated"

XX XX WO2004113365-A2.

XX XX 29-DEC-2004.

XX XX 19-MAY-2004; 2004WO-US015803.

XX XX 05-JUN-2003; 2003US-00454997.

XX XX 19-MAY-2004; 2004US-00849107.

XX XX (ENAN-) ENANTA PHARM INC.

XX PI Miao Z, Sun Y, Nakajima S, Tang D, Wang Z, Or YS;

XX XX WPI; 2005-091230/10.

XX DR New tripeptide compounds are potent hepatitis C virus replication
 XX PT inhibitors useful in the treatment of hepatitis C virus infection.

XX FS Disclosure; Page 24; 84pp; English.

XX CC The invention relates to a tripeptide compound (of formulae given in the
 CC specification) or its salt, ester or prodrug. Also included is treating
 CC hepatitis C viral infection in a subject involving administering a
 CC pharmaceutical composition comprising the tripeptide compound in
 CC combination with a carrier or excipient. The compounds are used for
 CC inhibiting replication of hepatitis C virus (by inhibition of e.g. NS3
 CC protease) in the treatment of hepatitis C virus infection and as
 CC laboratory tool to aid isolation of virally encoded serine protease. The
 CC compound is potent serine protease and HCV replication inhibitor and
 CC lessens the severity and adverse effects of HCV infection. The present
 CC sequence is a control HCV inhibitory peptide used in an assay to

CC determine inhibition of HCV by the compounds of the invention.
 XX SQ Sequence 6 AA;

Query Match 5.1%; Score 6; DB 9; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
 Db 1 DEMEEC 6
 |||||

RESULT 979

ADW97223
 ID ADW97223 standard; peptide; 6 AA.

XX AC ADW97223;

XX XX 21-APR-2005 (first entry)

XX DE HCV NS3 protease inhibitor Inh1.

XX KW Pharmaceutical; serine protease inhibitor; hepatitis C virus infection;
 KW antiinflammatory; hepatotropic; virucide; gastrointestinal disease;
 KW infection; antiviral.

XX OS Synthetic.

XX FH Key Location/Qualifiers
 XX FT Modified-site 1 /note= "Asp is acetylated"

XX XX WO2005010029-A1.

XX XX 03-FEB-2005.

XX XX 19-MAY-2004; 2004WO-US015802.

XX XX 03-JUL-2003; 2003US-00613206.

XX XX (ENAN-) ENANTA PHARM INC.

XX PI Wu FXH, Nakajima S, Or YS, Lu Z, Sun Y, Miao Z, Wang Z;

XX XX WPI; 2005-202245/21.

XX DR New aza-peptide macrocycles are hepatitis C virus replication inhibitor
 XX PT useful for the treatment of hepatitis C viral infections.

XX PS Example 25; Page 54; 75pp; English.

XX CC The invention relates to aza-peptide macrocycles (I) of formulae given in
 CC the specification. Also included are the preparation of (I) and a method
 CC of inhibiting the replication of hepatitis C virus comprising supplying a
 CC hepatitis C viral NS3 protease inhibitor amount of (I). The method of
 CC inhibiting the replication further comprises concurrent administration of
 CC an additional anti-hepatitis C virus agent such as alpha-interferon, beta
 CC -interferon, ribavirin or adamantane. The additional agent is an
 CC inhibitor of another target in the hepatitis C virus life cycle, which is
 CC helicase, polymerase, metalloprotease or internal ribosome entry site
 CC (IRES). The aza-peptide macrocycles are useful for the treatment of
 CC hepatitis C virus infections. The present sequence is an HCV NS3-NS4A
 CC protease inhibitory peptide used in an NS3-NS4A enzyme assay.

XX SQ Sequence 6 AA;

Query Match 5.1%; Score 6; DB 9; Length 6;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
 |||||

Db 1 DMEEC 6

RESULT 980
AEB25213
ID AEB25213 standard; peptide; 6 AA.
XX
AC AEB25213;
XX
XX 22-SEP-2005 (first entry)
XX
XX Peptide inhibitor HCV Inh 1, seq id 2.
XX
XX Antiviral; hepatitis C virus infection; inhibitor.
XX
XX Synthetic.
XX
XX US2005153877-A1.
XX
XX 14-JUL-2005.
XX
XX 06-FEB-2004; 2004US-00774047.
XX
XX 13-FEB-2003; 2003US-0509069P.
XX
XX (MIAO/) MIAO Z.
XX (SUNY/) SUN Y.
XX (NAKA/) NAKAJIMA S.
XX (TANG/) TANG D.
XX (WUFF/) WU F.
XX (XUGG/) XU G.
XX (ORIS/) OR Y S.
XX (WANG/) WANG Z.
XX
XX Miao Z, Sun Y, Nakajima S, Tang D, Wu F, Xu G, Or YS, Wang Z;
XX
XX WPI; 2005-531385/54.
XX
XX New macrocyclic compounds are hepatitis C viral non-structural protein-3
XX protease inhibitors, useful to treat hepatitis C viral infection.
XX
XX Example 215; SEQ ID NO 2; 229pp; English.
XX
XX The invention relates to novel macrocyclic compounds (I). Further
XX disclosed is a composition comprising an anti-hepatitis C virally
XX effective amount of (I) or their salts, esters, or prodrugs in
XX combination with a carrier or excipient. Also disclosed is a method of
XX inhibiting the replication of hepatitis C virus comprising supplying a
XX hepatitis C viral NS3 protease inhibitory amount of (I) and preparations
XX of (I). Compounds of the invention have virucide and hepatitis C viral
XX non-structural protein-3 protease inhibitor activity. (I) are useful to
XX treat hepatitis C viral infection and to inhibit the replication of
XX hepatitis C virus. The current sequence represents the peptide inhibitor
XX HCV Inh 1. This is used as a reference compound in an assay of HCV
XX protease activity and inhibition in an example of the invention.
XX
XX Sequence 6 AA;

Query Match 5.1%; Score 6; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DMEEC 65
Db 1 DMEEC 6

RESULT 981
AAR35974
ID AAR35974 standard; protein; 8 AA.
XX
AC AAR35974;
XX

DT 25-MAR-2003 (revised)
DT 24-MAY-1993 (first entry)
XX
DE Hepatitis C virus (HCV) epitope.
XX
KW Hepatitis; liver disease; HCV1; monoclonal antibody; epitope;
KW immobilised reagent; immunoassay; diagnosis; detection; treatment;
KW infection.
XX
OS Hepatitis C virus type 1.
XX
PN WO9300365-A2.
XX
PD 07-JAN-1993.
XX
PF 24-JUN-1992; 92WO-US005388.
XX
PR 24-JUN-1991; 91US-00722489.
XX
PA (CHIR) CHIRON CORP.
XX
PI Chien DY, Rutter W;
XX
DR WPI; 1993-036334/04.
XX
XX Polypeptide(s) comprising truncated hepatitis C virus sequences - for
XX detection, prevention and treatment of hepatitis C infection.
XX
PS Example A; Page 36; 80pp; English.
XX
CC This octamer was found to be immunoreactive with anti-HCV anti-sera. In
CC the epitope mapping experiment three different samples of anti-sera were
CC reacted with the peptide octamer, and then incubated with HRP-labelled
CC goat anti-human Ig antiserum, to enable detection of binding. This epitope
CC starts from amino acid 1655 of the HCV polyprotein. This was found to be
CC a weak epitope. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 8 AA;

Query Match 5.1%; Score 6; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 TSTWVL 16
Db 3 TSTWVL 8

RESULT 982
AAJ01075
ID AAJ01075 standard; peptide; 8 AA.
XX
AC AAJ01075;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1066.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
XX 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMUNE INC.
XX


```
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 143; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 8 AA;
    Query Match      5.1%; Score 6; DB 4; Length 8;
    Best Local Similarity 100.0%; Pred. No. 2e+06;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 11 TSTWVL 16
    DB 3 TSTWVL 8
RESULT 986
AAJ01667
ID AAJ01667 standard; peptide; 8 AA.
AC AAJ01667;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1658.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX OS Hepatitis C virus.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 117; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 8 AA;
    Query Match      5.1%; Score 6; DB 4; Length 8;
    Best Local Similarity 100.0%; Pred. No. 2e+06;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 11 TSTWVL 16
    DB 2 TSTWVL 7
RESULT 988
AAJ00566
ID AAJ00566 standard; peptide; 8 AA.
XX AC AAJ00566;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #557.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
```

```
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 143; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
XX SQ Sequence 8 AA;
    Query Match      5.1%; Score 6; DB 4; Length 8;
    Best Local Similarity 100.0%; Pred. No. 2e+06;
    Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 11 TSTWVL 16
    DB 3 TSTWVL 8
RESULT 986
AAJ01667
ID AAJ01667 standard; peptide; 8 AA.
AC AAJ01667;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1658.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX OS Hepatitis C virus.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX PT A new composition useful as a vaccines against hepatitis C virus.
XX PS Disclosure; Page 142; 214pp; English.
XX CC The present invention describes a composition comprising a prepared
XX CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX CC These are derived from HCV HLA-binding motifs. They are useful in
XX CC vaccines for the prevention and treatment of HCV infection in humans. The
XX CC present sequence is an epitope used in the disclosure of the invention
```

XX antiviral.
XX Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX 19-JUL-2000; 2000WO-US019774.
XX 19-JUL-1999; 99US-00357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 114; 214pp; English.
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX Sequence 8 AA;
XX Query Match 5.1%; Score 6; DB 4; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 4 SADLEV 9
XX DB 1 SADLEV 6
XX RESULT 989
XX AAJ01377
XX ID AAJ01377 standard; peptide; 8 AA.
XX AC AAJ01377;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1368.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX 19-JUL-2000; 2000WO-US019774.
XX 19-JUL-1999; 99US-00357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 135; 214pp; English.

XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX Sequence 8 AA;
XX Query Match 5.1%; Score 6; DB 4; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 11 TSTWVL 16
XX DB 2 TSTWVL 7
XX RESULT 990
XX AAJ02350
XX ID AAJ02350 standard; peptide; 8 AA.
XX AC AAJ02350;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #2341.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX antiviral.
XX Hepatitis C virus.
XX WO200121189-A1.
XX 29-MAR-2001.
XX 19-JUL-2000; 2000WO-US019774.
XX 19-JUL-1999; 99US-00357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccines against hepatitis C virus.
XX Disclosure; Page 158; 214pp; English.
XX The present invention describes a composition comprising a prepared
XX hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
XX These are derived from HCV HLA-binding motifs. They are useful in
XX vaccines for the prevention and treatment of HCV infection in humans. The
XX present sequence is an epitope used in the disclosure of the invention
XX Sequence 8 AA;
XX Query Match 5.1%; Score 6; DB 4; Length 8;
XX Best Local Similarity 100.0%; Pred. No. 2e+06;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 18 GGVLA 23
XX DB 3 GGVLA 8
XX RESULT 991
XX AAJ02350
XX ID AAJ02350 standard; peptide; 8 AA.
XX XX

CC AAY45390 to AAY48214 represent specifically claimed immunogenic peptides
 CC having a human major histocompatibility complex (MHC) Class I (also known
 CC as human leukocyte antigen (HLA)) binding motif. The immunogenic peptides
 CC can bind to a specific HLA allele (i.e. HLA-A subtypes HLA-A2.1, A1, A3.2
 CC or A24.1 or HLA-B or C) and induce a cytotoxic T cell response against
 CC the antigen from which the peptide is derived. Cytotoxic T lymphocytes
 CC (CTLs) which destroy antigen-bearing cells are normally induced by an
 CC antigen in the form of a peptide fragment bound to a HLA molecule, rather
 CC than the intact foreign antigen itself, and are particularly important in
 CC tumour rejection and in fighting viral infections. The peptides are
 CC therefore useful therapeutically to treat or prevent viral infections and
 CC cancers in mammals (especially humans) e.g. prostate cancer, hepatitis B
 CC and C, AIDS, and renal carcinoma. They can be administered as vaccines to
 CC elicit an immune response in individuals susceptible or otherwise at risk
 CC of viral infection or cancer, or used to treat chronic or acute
 CC conditions. They are also useful diagnostically, and can be used to
 CC induce a cytotoxic T cell response, by contacting a cytotoxic T cell with
 CC the peptide e.g. to produce CTLs ex vivo for infusion back into a
 CC patient. The polynucleotides encoding the immunogenic peptides are also
 CC useful therapeutically and for immunisation as above

XX Sequence 9 AA;

Query Match 5.1%; Score 6; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20

Db 1 VLLGGV 6
 |||||

RESULT 994

AAY73108
 ID AAY73108 standard; peptide; 9 AA.

XX AC AAY73108;

XX 06-AUG-2003 (revised)

DT 28-FEB-2000 (first entry)

XX Hepatitis C virus (HCV)-derived MHC class I (CTL) epitope, #266.

XX Chimeric; pan DR epitope; expression vector; promoter;
 KW major histocompatibility complex; MHC; targeting; peptide; epitope;
 KW antigen; presentation; class I; cytosolic pathway; endoplasmic reticulum;
 KW class II; extracellular antigen; endocytic pathway; helper T lymphocyte;
 KW HTL; universal epitope; cytotoxic T lymphocyte; CTL; immune response;
 KW immunogenicity; assay; vaccine; immunity; infection; pathogen; virus;
 KW HIV; HBV; HCV; hepatitis B; hepatitis C; bacterium; protozoan;
 KW tumour cell; autoimmune disease; activation; antiviral; antimalarial;
 KW immunoprotective.

XX Synthetic.

OS Hepatitis C virus.

XX WO9958658-A2.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-US010646.

XX 13-MAY-1998; 98US-00078904.

XX 15-MAY-1998; 98US-0085751P.

XX (EPIM-) EPIMUNE INC.

XX Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;
 XX Cheenut RW;

XX WPI; 2000-039103/03.

XX Expression vectors encoding major histocompatibility targeting sequence,

PT used as, e.g. tumor vaccines.

XX Claim 11; Page 68; 130pp; English.

XX Sequences AAY73103-V73145 represent hepatitis C virus (HCV)-derived MHC
 CC class I (CTL) epitopes which are claimed for use in the present
 CC invention. The invention relates to a novel expression vector comprising
 CC a promoter operably linked to a fusion gene encoding a major
 CC histocompatibility complex (MHC) targeting sequence, and two or more
 CC heterologous peptide epitopes. The MHC targeting sequence may be a class
 CC I targeting sequence, which directs an MHC class I epitope to a
 CC cytosolic pathway or to the endoplasmic reticulum, or an MHC class II
 CC targeting sequence, which directs extracellular antigens to enter the
 CC endocytic pathway to be processed into antigen peptides for presentation
 CC on MHC class II molecules. The heterologous epitopes may comprise either
 CC helper T lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL)
 CC epitope and a universal HTL epitope such as a pan DR epitope (PADRE). The
 CC vectors are useful for stimulating an immune response in vivo, as well as
 CC for use in assaying the human immunogenicity of a human T cell peptide
 CC epitope in vivo in a non-human mammal. They provide a nucleic acid
 CC vaccine for enhancing immunity against infectious pathogens, such as
 CC viruses (e.g., HIV, hepatitis B (HBV) and hepatitis C (HCV)), bacteria,
 CC protozoa (e.g., Plasmodium falciparum, the cause of malaria) and also
 CC tumour cells and autoimmune diseases. Universal MHC class II epitopes are
 CC advantageously combined with other MHC class I and class II epitopes to
 CC increase the number of cells that are activated in response to a given
 CC antigen and provide a broader population coverage of MHC-reactive
 CC alleles. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 9 AA;

Query Match 5.1%; Score 6; DB 3; Length 9;

Best Local Similarity 100.0%; Pred. No. 2e+06;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23

Db 4 GGVLA 9
 |||||

RESULT 995

AAY73304

ID AAY73304 standard; peptide; 9 AA.

XX AC AAY73304;

XX 28-FEB-2000 (first entry)

XX Plasmodium falciparum-derived MHC class I (CTL) epitope, #462.

XX Chimeric; pan DR epitope; expression vector; promoter;
 KW major histocompatibility complex; MHC; targeting; peptide; epitope;
 KW antigen; presentation; class I; cytosolic pathway; endoplasmic reticulum;
 KW class II; extracellular antigen; endocytic pathway; helper T lymphocyte;
 KW HTL; universal epitope; cytotoxic T lymphocyte; CTL; immune response;
 KW immunogenicity; assay; vaccine; immunity; infection; pathogen; virus;
 KW HIV; HBV; HCV; hepatitis B; hepatitis C; bacterium; protozoan;
 KW tumour cell; autoimmune disease; activation; antiviral; antimalarial;
 KW immunoprotective.

XX Synthetic.

OS Plasmodium falciparum.

XX WO9958658-A2.

XX 18-NOV-1999.

XX 13-MAY-1999; 99WO-US010646.

XX 13-MAY-1998; 98US-00078904.

XX 15-MAY-1998; 98US-0085751P.

XX (EPIM-) EPIMUNE INC.

XX Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;
 PI Chesnut RW;
 XX WPI; 2000-039103/03.
 XX Expression vectors encoding major histocompatibility targeting sequence,
 PT used as, e.g. tumor vaccines.
 XX Claim 11; Page 75; 130pp; English.
 XX Sequences AAU73283-Y73314 represent Plasmidium falciparum-derived MHC
 CC class I (CTL) epitopes which are claimed for use in the present
 CC invention. The invention relates to a novel expression vector comprising
 CC a promoter operably linked to a fusion gene encoding a major
 CC histocompatibility complex (MHC) targeting sequence, and two or more
 CC heterologous peptide epitopes. The MHC targeting sequence may be a class
 CC I targeting sequence, which directs an MHC class I epitope to a
 CC cytosolic pathway or to the endoplasmic reticulum, or an MHC class II
 CC targeting sequence, which directs extracellular antigens to enter the
 CC endocytic pathway to be processed into antigen peptides for presentation
 CC on MHC class II molecules. The heterologous epitopes may comprise either
 CC helper T lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL)
 CC epitope and a universal HTL epitope such as a pan DR epitope (PADRG). The
 CC vectors are useful for stimulating an immune response in vivo, as well as
 CC for use in assaying the human immunogenicity of a human T cell peptide
 CC epitope in vivo in a non-human mammal. They provide a nucleic acid
 CC vaccine for enhancing immunity against infectious pathogens, such as
 CC viruses (e.g., HIV, hepatitis B (HBV) and hepatitis C (HCV)) bacteria,
 CC protozoa (e.g., Plasmidium falciparum, the cause of malaria) and also
 CC tumour cells and autoimmune diseases. Universal MHC class II epitopes are
 CC advantageously combined with other MHC class I and class II epitopes to
 CC increase the number of cells that are activated in response to a given
 CC antigen and provide a broader population coverage of MHC-reactive alleles
 XX Sequence 9 AA;
 SQ

Query Match 5.1%; Score 6; DB 3; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
 Db 1 VLLGGV 6
 |||||
 |||||

RESULT 996
 AAU26982
 ID AAU26982 standard; peptide; 9 AA.
 XX AC AAU26982;
 XX 18-DEC-2001 (first entry)
 XX Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #266.
 XX Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
 KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
 KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
 KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
 XX Homo sapiens.
 XX OS
 XX WO200162776-A1.
 XX 30-AUG-2001.
 XX 23-FEB-2000; 2000WO-US004655.
 XX 23-FEB-2000; 2000WO-US004655.
 XX (EPIM-) EPIMUNE INC.
 XX Sette A, Sidney J, Kast WM, Southwood S;
 XX WPI; 2001-582039/65.
 XX Composition for treating viral diseases and cancer comprises an
 PT immunogenic peptide having an HLA-A2.1 binding motif.
 XX Example 1; Page 30; 85pp; English.
 XX Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
 CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
 CC the invention are capable of specifically binding glycoproteins encoded
 CC by HLA alleles and inducing a cytotoxic T cell response against an
 CC antigen in a patient expressing HLA-A2.1. This method is useful for the
 CC treatment, prevention and diagnosis of pathological states such as viral
 CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
 CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
 CC acuminatum. The peptides are used for treatment of chronic infection and
 CC for stimulating the immune system to eliminate virus-infected cells
 XX Sequence 9 AA;
 SQ

PI Sette A, Sidney J, Kast WM, Southwood S;
 XX WPI; 2001-582039/65.
 XX Composition for treating viral diseases and cancer comprises an
 PT immunogenic peptide having an HLA-A2.1 binding motif.
 XX Claim 1; Page 68; 85pp; English.
 XX Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
 CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
 CC the invention are capable of specifically binding glycoproteins encoded
 CC by HLA alleles and inducing a cytotoxic T cell response against an
 CC antigen in a patient expressing HLA-A2.1. This method is useful for the
 CC treatment, prevention and diagnosis of pathological states such as viral
 CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
 CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
 CC acuminatum. The peptides are used for treatment of chronic infection and
 CC for stimulating the immune system to eliminate virus-infected cells
 XX Sequence 9 AA;
 SQ

Query Match 5.1%; Score 6; DB 4; Length 9;
 Best Local Similarity 100.0%; Pred. No. 2e+06;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADL 7
 Db 4 CMSADL 9
 |||||
 |||||

RESULT 997
 AAU26649
 ID AAU26649 standard; peptide; 9 AA.
 XX AC AAU26649;
 XX 18-DEC-2001 (first entry)
 XX Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #92.
 XX Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
 KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
 KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
 KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
 XX Homo sapiens.
 XX OS
 XX WO200162776-A1.
 XX 30-AUG-2001.
 XX 23-FEB-2000; 2000WO-US004655.
 XX 23-FEB-2000; 2000WO-US004655.
 XX (EPIM-) EPIMUNE INC.
 XX Sette A, Sidney J, Kast WM, Southwood S;
 XX WPI; 2001-582039/65.
 XX Composition for treating viral diseases and cancer comprises an
 PT immunogenic peptide having an HLA-A2.1 binding motif.
 XX Example 1; Page 30; 85pp; English.
 XX Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
 CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
 CC the invention are capable of specifically binding glycoproteins encoded
 CC by HLA alleles and inducing a cytotoxic T cell response against an
 CC antigen in a patient expressing HLA-A2.1. This method is useful for the
 CC treatment, prevention and diagnosis of pathological states such as viral

CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma acuminatum. The peptides are used for treatment of chronic infection and for stimulating the immune system to eliminate virus-infected cells

XX Sequence 9 AA;

Query Match 5.1%; Score 6; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 CWSADL 7
|||||

Db 4 CWSADL 9
|||||

RESULT 998
AAJ02575
ID AAJ02575 standard; peptide; 9 AA.
XX
AC AAJ02575;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #2566.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
FN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 163; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121. These are derived from HCV HLA-binding motifs. They are useful in CC vaccines for the prevention and treatment of HCV infection in humans. The CC present sequence is an epitope used in the disclosure of the invention

XX Sequence 9 AA;

Query Match 5.1%; Score 6; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 GGVLA 23
|||||

Db 4 GGVLA 9
|||||

RESULT 999
AAJ01832
ID AAJ01832 standard; peptide; 9 AA.
XX
AC AAJ01832;
XX

DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1823.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
FN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 146; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121. These are derived from HCV HLA-binding motifs. They are useful in CC vaccines for the prevention and treatment of HCV infection in humans. The CC present sequence is an epitope used in the disclosure of the invention

XX Sequence 9 AA;

Query Match 5.1%; Score 6; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 GGVLA 23
|||||

Db 4 GGVLA 9
|||||

RESULT 1000
AAJ03208
ID AAJ03208 standard; peptide; 9 AA.
XX
AC AAJ03208;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3199.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX
OS Hepatitis C virus.
XX
FN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX

```

DR  WPI; 2001-308046/32.
XX
PT  A new composition useful as a vaccines against hepatitis C virus.
XX
PS  Disclosure; Page 176; 214pp; English.
XX
CC  The present invention describes a composition comprising a prepared
CC  hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC  These are derived from HCV HLA-binding motifs. They are useful in
CC  vaccines for the prevention and treatment of HCV infection in humans. The
CC  present sequence is an epitope used in the disclosure of the invention
XX
SQ  Sequence 9 AA;

Query Match      5.1%; Score 6; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  18 GGVLA 23
    |||||
Db   4 GGVLA 9

Search completed: January 27, 2006, 19:29:23
Job time : 199 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:22:10 ; Search time 23 Seconds
(without alignments)
424.162 Million cell updates/sec

Title: US-09-638-693A-36_COPY_16_133

Perfect score: 118

Sequence: 1 ACSADLEVTSTWLLGGV.....VIEPIVTTNMQKLEAFMHKH 118

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 572060 seqs, 82675679 residues

Word size: 3, 0, 0, 0

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 2000 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	118	100.0	133	2	US-09-878-281A-36
2	83	70.3	133	2	US-09-878-281A-38
3	83	70.3	133	2	US-09-878-281A-40
4	83	70.3	209	2	US-09-878-281A-223
5	47	39.8	128	1	US-08-244-116B-17
6	47	39.8	133	2	US-09-878-281A-32
7	47	39.8	133	2	US-09-878-281A-34
8	41	34.7	829	2	US-09-881-239-5
9	41	34.7	1099	2	US-10-637-323-4
10	41	34.7	1099	2	US-10-637-323-4
11	20	16.9	20	2	US-09-790-497A-104
12	20	16.9	20	2	US-09-878-281A-97
13	20	16.9	22	1	US-08-146-028-107
14	20	16.9	22	2	US-08-723-425A-107
15	20	16.9	22	2	US-09-112-206-107
16	20	16.9	22	2	US-09-576-824A-104
17	20	16.9	22	2	US-09-680-497-107
18	19	16.1	19	1	US-08-244-116B-2
19	19	16.1	19	2	US-08-537-802-44
20	18	15.3	20	2	US-09-790-497A-105
21	18	15.3	20	2	US-09-878-281A-99
22	18	15.3	22	1	US-08-146-028-108
23	18	15.3	22	2	US-08-723-425A-108
24	18	15.3	22	2	US-09-112-206-108
25	18	15.3	22	2	US-09-576-824A-105
26	18	15.3	22	2	US-09-680-497-108
27	18	15.3	352	2	US-08-921-887-52
28	18	15.3	18	13	US-09-878-281A-36
29	14	11.9	18	13	US-09-878-281A-38
30	14	11.9	18	13	US-09-878-281A-40
31	13	11.0	20	2	US-09-878-281A-223
32	13	11.0	128	1	US-08-244-116B-17
33	13	11.0	133	2	US-09-878-281A-32
34	13	11.0	133	2	US-09-878-281A-34
35	13	11.0	829	2	US-09-881-239-5
36	13	11.0	1099	2	US-10-637-323-4
37	13	11.0	1099	2	US-10-637-323-4
38	13	11.0	20	2	US-09-790-497A-104
39	13	11.0	20	2	US-09-878-281A-97
40	13	11.0	22	1	US-08-146-028-107
41	13	11.0	22	2	US-08-723-425A-107
42	13	11.0	22	2	US-09-112-206-107
43	13	11.0	22	2	US-09-576-824A-104
44	13	11.0	22	2	US-09-680-497-107
45	13	11.0	19	1	US-08-244-116B-2
46	13	11.0	19	2	US-08-537-802-44
47	13	11.0	20	2	US-09-790-497A-105
48	13	11.0	20	2	US-09-878-281A-99
49	13	11.0	22	1	US-08-146-028-108
50	13	11.0	22	2	US-08-723-425A-108
51	13	11.0	22	2	US-09-112-206-108
52	13	11.0	22	2	US-09-576-824A-105
53	13	11.0	22	2	US-09-680-497-108
54	13	11.0	352	2	US-08-921-887-52
55	13	11.0	18	13	US-09-878-281A-36
56	13	11.0	18	13	US-09-878-281A-38
57	13	11.0	18	13	US-09-878-281A-40
58	13	11.0	20	2	US-09-878-281A-223
59	13	11.0	128	1	US-08-244-116B-17
60	13	11.0	133	2	US-09-878-281A-32
61	13	11.0	133	2	US-09-878-281A-34
62	13	11.0	829	2	US-09-881-239-5
63	13	11.0	1099	2	US-10-637-323-4
64	13	11.0	1099	2	US-10-637-323-4
65	13	11.0	20	2	US-09-790-497A-104
66	13	11.0	20	2	US-09-878-281A-97
67	13	11.0	22	1	US-08-146-028-107
68	13	11.0	22	2	US-08-723-425A-107
69	13	11.0	22	2	US-09-112-206-107
70	13	11.0	22	2	US-09-576-824A-104
71	13	11.0	22	2	US-09-680-497-107
72	13	11.0	19	1	US-08-244-116B-2
73	13	11.0	19	2	US-08-537-802-44
74	13	11.0	20	2	US-09-790-497A-105
75	13	11.0	20	2	US-09-878-281A-99
76	13	11.0	22	1	US-08-146-028-108
77	13	11.0	22	2	US-08-723-425A-108
78	13	11.0	22	2	US-09-112-206-108
79	13	11.0	22	2	US-09-576-824A-105
80	13	11.0	22	2	US-09-680-497-108
81	13	11.0	352	2	US-08-921-887-52
82	13	11.0	18	13	US-09-878-281A-36
83	13	11.0	18	13	US-09-878-281A-38
84	13	11.0	18	13	US-09-878-281A-40
85	13	11.0	20	2	US-09-878-281A-223
86	13	11.0	128	1	US-08-244-116B-17
87	13	11.0	133	2	US-09-878-281A-32
88	13	11.0	133	2	US-09-878-281A-34
89	13	11.0	829	2	US-09-881-239-5
90	13	11.0	1099	2	US-10-637-323-4
91	13	11.0	1099	2	US-10-637-323-4
92	13	11.0	20	2	US-09-790-497A-104
93	13	11.0	20	2	US-09-878-281A-97
94	13	11.0	22	1	US-08-146-028-107
95	13	11.0	22	2	US-08-723-425A-107
96	13	11.0	22	2	US-09-112-206-107
97	13	11.0	22	2	US-09-576-824A-104
98	13	11.0	22	2	US-09-680-497-107
99	13	11.0	19	1	US-08-244-116B-2
100	13	11.0	19	2	US-08-537-802-44

Sequence 52, Appl
Sequence 1, Appl
Sequence 30, Appl
Sequence 98, Appl
Sequence 100, Appl
Sequence 109, App
Sequence 106, App
Sequence 109, App
Sequence 106, App
Sequence 109, App
Sequence 26, Appl
Sequence 7, Appl
Sequence 30, Appl
Sequence 13, Appl
Sequence 12, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 14, Appl
Sequence 7, Appl
Sequence 45, Appl
Sequence 7, Appl
Sequence 45, Appl
Sequence 7, Appl
Sequence 45, Appl
Sequence 7, Appl
Sequence 45, Appl
Sequence 40, Appl
Sequence 4, Appl
Sequence 41, Appl
Sequence 41, Appl
Sequence 43, Appl
Sequence 43, Appl
Sequence 44, Appl
Sequence 44, Appl
Sequence 42, Appl
Sequence 42, Appl
Sequence 68, Appl
Sequence 20, Appl
Sequence 47, Appl
Sequence 47, Appl
Sequence 48, Appl
Sequence 48, Appl
Sequence 16, Appl
Sequence 21, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 23, Appl
Sequence 49, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 17, Appl
Sequence 31, Appl
Sequence 32, Appl
Sequence 43, Appl
Sequence 43, Appl
Sequence 45, Appl
Sequence 46, Appl
Sequence 47, Appl
Sequence 48, Appl
Sequence 49, Appl

101	13	11.0	686	2	US-09-930-591-2	Sequence 2, Appli	174	12	10.2	54	1	US-08-936-865-2	Sequence 2, Appli
102	13	11.0	686	2	US-09-930-591-3	Sequence 3, Appli	175	12	10.2	54	2	US-09-198-723A-24	Sequence 24, Appl
103	13	11.0	686	2	US-09-930-591-4	Sequence 4, Appli	176	12	10.2	54	2	US-09-684-881-24	Sequence 24, Appl
104	13	11.0	686	2	US-09-930-591-5	Sequence 5, Appli	177	12	10.2	247	1	US-08-324-977-44	Sequence 44, Appl
105	13	11.0	686	2	US-09-930-591-6	Sequence 6, Appli	178	12	10.2	247	1	US-08-384-616-44	Sequence 44, Appl
106	13	11.0	686	2	US-09-930-591-7	Sequence 7, Appli	179	12	10.2	247	1	US-08-904-686A-44	Sequence 44, Appl
107	13	11.0	686	2	US-09-930-591-8	Sequence 8, Appli	180	12	10.2	247	2	US-09-315-850-44	Sequence 44, Appl
108	13	11.0	686	2	US-09-930-591-9	Sequence 9, Appli	181	12	10.2	675	1	US-08-371-036-2	Sequence 2, Appli
109	13	11.0	686	2	US-09-930-591-10	Sequence 10, Appl	182	12	10.2	675	2	US-09-096-570-2	Sequence 2, Appli
110	13	11.0	686	2	US-09-930-591-11	Sequence 11, Appl	183	12	10.2	675	2	US-09-365-617B-2	Sequence 2, Appli
111	13	11.0	728	2	US-09-881-239-1	Sequence 11, Appl	184	12	10.2	1692	2	US-09-263-933-4	Sequence 4, Appli
112	13	11.0	739	2	US-08-444-818-148	Sequence 148, App	185	12	10.2	1692	2	US-09-263-933-11	Sequence 11, Appl
113	13	11.0	859	2	US-08-444-818-30	Sequence 30, Appl	186	12	10.2	1692	2	US-09-263-933-18	Sequence 18, Appl
114	13	11.0	971	2	US-08-867-611-52	Sequence 52, Appl	187	12	10.2	1692	2	US-09-919-901-4	Sequence 4, Appli
115	13	11.0	971	2	US-09-690-359-52	Sequence 52, Appl	188	12	10.2	1692	2	US-09-919-901-11	Sequence 11, Appl
116	13	11.0	973	2	US-08-867-611-53	Sequence 53, Appl	189	12	10.2	1692	2	US-09-919-901-18	Sequence 18, Appl
117	13	11.0	973	2	US-09-690-359-53	Sequence 53, Appl	190	12	10.2	1692	2	US-10-191-966-4	Sequence 4, Appli
118	13	11.0	992	2	US-08-867-611-54	Sequence 54, Appl	191	12	10.2	1692	2	US-10-191-966-11	Sequence 11, Appl
119	13	11.0	992	2	US-09-690-359-54	Sequence 54, Appl	192	12	10.2	1692	2	US-10-191-966-18	Sequence 18, Appl
120	13	11.0	1021	1	US-07-910-760-12	Sequence 12, Appl	193	12	10.2	1985	2	US-09-539-601-9	Sequence 9, Appli
121	13	11.0	1021	1	US-08-440-519-12	Sequence 12, Appl	194	12	10.2	1985	2	US-09-539-601-12	Sequence 12, Appl
122	13	11.0	1021	2	US-08-440-549-12	Sequence 12, Appl	195	12	10.2	1985	2	US-09-539-601-18	Sequence 18, Appl
123	13	11.0	2261	2	US-08-444-818-66	Sequence 66, Appl	196	12	10.2	1985	2	US-09-539-601-24	Sequence 24, Appl
124	13	11.0	2436	2	US-08-444-818-75	Sequence 75, Appl	197	12	10.2	1985	2	US-09-539-601-30	Sequence 30, Appl
125	13	11.0	2772	2	US-08-444-818-89	Sequence 89, Appl	198	12	10.2	1985	2	US-10-359-275-42	Sequence 42, Appl
126	13	11.0	2894	1	US-08-466-975A-23	Sequence 23, Appl	199	12	10.2	2013	1	US-08-324-977-12	Sequence 12, Appl
127	13	11.0	2894	1	US-08-391-671A-23	Sequence 23, Appl	200	12	10.2	2013	1	US-08-384-616-12	Sequence 12, Appl
128	13	11.0	2894	2	US-08-467-902A-23	Sequence 23, Appl	201	12	10.2	2013	1	US-08-904-686A-12	Sequence 12, Appl
129	13	11.0	2894	2	US-09-275-265-23	Sequence 23, Appl	202	12	10.2	2013	2	US-09-315-850-12	Sequence 12, Appl
130	13	11.0	2894	2	US-09-941-611-23	Sequence 23, Appl	203	12	10.2	2201	2	US-08-952-981A-2	Sequence 2, Appli
131	13	11.0	2894	2	US-10-044-995-23	Sequence 23, Appl	204	12	10.2	2201	2	US-09-539-601-6	Sequence 6, Appli
132	13	11.0	2955	1	US-08-443-260-3	Sequence 3, Appli	205	12	10.2	2201	2	US-09-539-601-15	Sequence 15, Appl
133	13	11.0	2955	2	US-08-442-805A-3	Sequence 3, Appli	206	12	10.2	2201	2	US-10-029-907-3	Sequence 3, Appli
134	13	11.0	2955	2	US-08-443-900A-3	Sequence 3, Appli	207	12	10.2	2201	2	US-10-309-561A-3	Sequence 3, Appli
135	13	11.0	2955	2	US-08-444-818-124	Sequence 124, App	208	12	10.2	2307	2	US-09-263-933-2	Sequence 2, Appli
136	13	11.0	2955	2	US-08-249-843-3	Sequence 3, Appli	209	12	10.2	2307	2	US-09-263-933-9	Sequence 9, Appli
137	13	11.0	2995	2	US-08-444-818-138	Sequence 138, App	210	12	10.2	2307	2	US-09-263-933-16	Sequence 16, Appl
138	13	11.0	3011	1	US-08-188-281B-1	Sequence 1, Appli	211	12	10.2	2307	2	US-09-919-901-2	Sequence 2, Appli
139	13	11.0	3011	1	US-08-453-552-1	Sequence 1, Appli	212	12	10.2	2307	2	US-09-919-901-9	Sequence 9, Appli
140	13	11.0	3011	1	US-08-453-552-2	Sequence 2, Appli	213	12	10.2	2307	2	US-09-919-901-16	Sequence 16, Appl
141	13	11.0	3011	1	US-08-440-103-36	Sequence 36, Appl	214	12	10.2	2307	2	US-10-191-966-2	Sequence 2, Appli
142	13	11.0	3011	1	US-08-440-542-36	Sequence 36, Appl	215	12	10.2	2307	2	US-10-191-966-9	Sequence 9, Appli
143	13	11.0	3011	1	US-07-910-760-10	Sequence 10, Appl	216	12	10.2	2307	2	US-10-191-966-16	Sequence 16, Appl
144	13	11.0	3011	1	US-08-440-519-10	Sequence 10, Appl	217	12	10.2	2620	1	US-08-324-977-32	Sequence 32, Appl
145	13	11.0	3011	1	US-08-231-368-36	Sequence 36, Appl	218	12	10.2	2620	1	US-08-384-616-32	Sequence 32, Appl
146	13	11.0	3011	1	US-08-440-210-36	Sequence 36, Appl	219	12	10.2	2620	1	US-08-904-686A-32	Sequence 32, Appl
147	13	11.0	3011	1	US-08-710-637-1	Sequence 1, Appli	220	12	10.2	2620	2	US-09-315-850-32	Sequence 32, Appl
148	13	11.0	3011	1	US-08-710-637-2	Sequence 2, Appli	221	12	10.2	2621	1	US-08-324-977-36	Sequence 36, Appl
149	13	11.0	3011	1	US-08-833-678A-6	Sequence 6, Appli	222	12	10.2	2621	1	US-08-384-616-36	Sequence 36, Appl
150	13	11.0	3011	2	US-08-811-566-20	Sequence 20, Appl	223	12	10.2	2621	1	US-08-904-686A-36	Sequence 36, Appl
151	13	11.0	3011	2	US-08-444-818-177	Sequence 177, App	224	12	10.2	2621	2	US-09-315-850-36	Sequence 36, Appl
152	13	11.0	3011	2	US-09-014-416-1	Sequence 1, Appli	225	12	10.2	2985	2	US-10-259-275-40	Sequence 40, Appl
153	13	11.0	3011	2	US-09-014-416-5	Sequence 5, Appli	226	12	10.2	3010	1	US-08-324-977-2	Sequence 2, Appli
154	13	11.0	3011	2	US-08-529-169A-6	Sequence 6, Appli	227	12	10.2	3010	1	US-08-324-977-14	Sequence 14, Appl
155	13	11.0	3011	2	US-09-388-874-2	Sequence 2, Appli	228	12	10.2	3010	1	US-08-384-616-2	Sequence 2, Appli
156	13	11.0	3011	2	US-08-046-604-36	Sequence 36, Appl	229	12	10.2	3010	1	US-08-384-616-14	Sequence 14, Appl
157	13	11.0	3011	2	US-08-440-549-10	Sequence 10, Appl	230	12	10.2	3010	1	US-08-904-686A-2	Sequence 2, Appli
158	13	11.0	3011	2	US-08-850-328-1	Sequence 1, Appli	231	12	10.2	3010	1	US-08-904-686A-14	Sequence 14, Appl
159	13	11.0	3011	2	US-08-034-756-20	Sequence 20, Appl	232	12	10.2	3010	2	US-09-014-416-3	Sequence 3, Appli
160	13	11.0	3011	2	US-09-483-799-6	Sequence 6, Appli	233	12	10.2	3010	2	US-09-315-850-2	Sequence 2, Appli
161	13	11.0	3011	2	US-09-916-359-2	Sequence 2, Appli	234	12	10.2	3010	2	US-09-315-850-14	Sequence 14, Appl
162	13	11.0	3011	2	US-10-104-966-1	Sequence 1, Appli	235	12	10.2	3010	2	US-09-539-601-3	Sequence 3, Appli
163	13	11.0	3011	2	US-09-952-572-9	Sequence 9, Appli	236	12	10.2	3010	2	US-09-539-601-21	Sequence 21, Appl
164	13	11.0	3011	2	US-09-929-955-1	Sequence 1, Appli	237	12	10.2	3010	2	US-09-539-601-27	Sequence 27, Appl
165	13	11.0	3011	4	PCT-US91-02225-10	Sequence 20, Appl	238	12	10.2	3010	2	US-09-539-601-33	Sequence 33, Appl
166	13	11.0	3011	4	PCT-US91-02225-10	Sequence 20, Appl	239	11	9.3	20	2	US-09-878-281A-101	Sequence 101, App
167	13	11.0	3011	4	PCT-US93-00907-1	Sequence 1, Appli	240	10	8.5	10	2	US-08-197-484-60	Sequence 60, Appl
168	13	11.0	3011	4	PCT-US94-07280-2	Sequence 2, Appli	241	10	8.5	10	2	US-08-197-484-139	Sequence 139, App
169	13	11.0	3011	4	PCT-US94-07280-1	Sequence 1, Appli	242	10	8.5	10	4	PCT-US95-02121-60	Sequence 60, Appl
170	13	11.0	3011	4	PCT-US95-01087-1	Sequence 1, Appli	243	10	8.5	10	4	PCT-US95-02121-139	Sequence 139, App
171	13	11.0	3012	2	US-08-811-566-2	Sequence 2, Appli	244	10	8.5	15	2	US-09-009-953-256	Sequence 256, App
172	13	11.0	3012	2	US-09-034-756-2	Sequence 2, Appli	245	9	7.6	128	2	US-09-878-281A-62	Sequence 62, Appl
173	12	10.2	54	1	US-08-700-356-2	Sequence 2, Appli	246	9	7.6	481	2	US-09-878-281A-270	Sequence 270, App

247	9	7.6	484	2	US-09-878-281A-198	Sequence 198, App	320	8	6.8	20	2	US-09-275-265-10	Sequence 10, Appl
248	9	7.6	484	2	US-09-878-281A-200	Sequence 200, App	321	8	6.8	20	2	US-09-275-265-11	Sequence 11, Appl
249	9	7.6	631	1	US-08-700-356-1	Sequence 1, Appli	322	8	6.8	20	2	US-08-850-328-8	Sequence 8, Appl
250	9	7.6	631	1	US-08-936-865-1	Sequence 1, Appli	323	8	6.8	20	2	US-09-941-611-10	Sequence 10, Appl
251	9	7.6	632	2	US-09-198-723A-23	Sequence 23, Appl	324	8	6.8	20	2	US-09-941-611-11	Sequence 11, Appl
252	9	7.6	632	2	US-09-684-881-23	Sequence 23, Appl	325	8	6.8	20	2	US-08-775-052A-47	Sequence 47, Appl
253	9	7.6	646	2	US-09-198-723A-60	Sequence 60, Appl	326	8	6.8	20	2	US-09-790-497A-50	Sequence 50, Appl
254	9	7.6	646	2	US-09-198-723A-63	Sequence 63, Appl	327	8	6.8	20	2	US-09-790-497A-51	Sequence 51, Appl
255	9	7.6	646	2	US-09-198-723A-66	Sequence 66, Appl	328	8	6.8	20	2	US-09-790-497A-52	Sequence 52, Appl
256	9	7.6	646	2	US-09-198-723A-69	Sequence 69, Appl	329	8	6.8	20	2	US-09-790-497A-121	Sequence 121, App
257	9	7.6	646	2	US-09-198-723A-72	Sequence 72, Appl	330	8	6.8	20	2	US-09-576-824A-121	Sequence 121, App
258	9	7.6	646	2	US-09-684-881-60	Sequence 60, Appl	331	8	6.8	20	2	US-10-044-995-10	Sequence 10, Appl
259	9	7.6	646	2	US-09-684-881-63	Sequence 63, Appl	332	8	6.8	20	2	US-10-044-995-11	Sequence 11, Appl
260	9	7.6	646	2	US-09-684-881-66	Sequence 66, Appl	333	8	6.8	20	4	PCT-US94-05407-11	Sequence 11, Appl
261	9	7.6	646	2	US-09-684-881-69	Sequence 69, Appl	334	8	6.8	22	1	US-08-146-028-50	Sequence 50, Appl
262	9	7.6	646	2	US-09-684-881-72	Sequence 72, Appl	335	8	6.8	22	1	US-08-146-028-51	Sequence 51, Appl
263	9	7.6	665	2	US-09-543-376B-1	Sequence 1, Appli	336	8	6.8	22	1	US-08-146-028-52	Sequence 52, Appl
264	9	7.6	665	2	US-09-543-376B-2	Sequence 2, Appli	337	8	6.8	22	1	US-08-146-028-121	Sequence 121, App
265	9	7.6	665	2	US-09-543-376B-3	Sequence 3, Appli	338	8	6.8	22	2	US-08-723-425A-50	Sequence 50, Appl
266	9	7.6	666	2	US-09-198-723A-11	Sequence 11, Appl	339	8	6.8	22	2	US-08-723-425A-51	Sequence 51, Appl
267	9	7.6	666	2	US-09-198-723A-12	Sequence 12, Appl	340	8	6.8	22	2	US-08-723-425A-52	Sequence 52, Appl
268	9	7.6	666	2	US-09-198-723A-13	Sequence 13, Appl	341	8	6.8	22	2	US-08-723-425A-121	Sequence 121, App
269	9	7.6	666	2	US-09-198-723A-14	Sequence 14, Appl	342	8	6.8	22	2	US-09-112-206-50	Sequence 50, Appl
270	9	7.6	666	2	US-09-198-723A-15	Sequence 15, Appl	343	8	6.8	22	2	US-09-112-206-51	Sequence 51, Appl
271	9	7.6	666	2	US-09-198-723A-16	Sequence 16, Appl	344	8	6.8	22	2	US-09-112-206-52	Sequence 52, Appl
272	9	7.6	666	2	US-09-198-723A-17	Sequence 17, Appl	345	8	6.8	22	2	US-09-576-824A-50	Sequence 50, Appl
273	9	7.6	666	2	US-09-198-723A-18	Sequence 18, Appl	346	8	6.8	22	2	US-09-576-824A-51	Sequence 51, Appl
274	9	7.6	666	2	US-09-684-881-12	Sequence 12, Appl	347	8	6.8	22	2	US-09-680-497-52	Sequence 52, Appl
275	9	7.6	666	2	US-09-684-881-13	Sequence 13, Appl	348	8	6.8	22	2	US-09-680-497-121	Sequence 121, App
276	9	7.6	666	2	US-09-684-881-14	Sequence 14, Appl	349	8	6.8	25	2	PCT-US92-07865-1	Sequence 1, Appli
277	9	7.6	666	2	US-09-684-881-15	Sequence 15, Appl	350	8	6.8	25	2	US-09-929-955-25	Sequence 25, Appl
278	9	7.6	666	2	US-09-684-881-16	Sequence 16, Appl	351	8	6.8	25	2	US-09-929-955-26	Sequence 26, Appl
279	9	7.6	666	2	US-09-684-881-17	Sequence 17, Appl	352	8	6.8	25	2	US-09-929-955-27	Sequence 27, Appl
280	9	7.6	666	2	US-09-684-881-18	Sequence 18, Appl	353	8	6.8	25	2	US-09-929-955-33	Sequence 33, Appl
281	9	7.6	672	2	US-09-684-881-19	Sequence 19, Appl	354	8	6.8	25	2	US-09-929-955-34	Sequence 34, Appl
282	9	7.6	672	2	US-09-684-881-20	Sequence 20, Appl	355	8	6.8	25	2	US-09-929-955-35	Sequence 35, Appl
283	9	7.6	672	2	US-09-684-881-21	Sequence 21, Appl	356	8	6.8	25	2	US-09-929-955-36	Sequence 36, Appl
284	9	7.6	672	2	US-09-684-881-22	Sequence 22, App	357	8	6.8	25	2	US-09-929-955-37	Sequence 37, Appl
285	9	7.6	672	2	US-09-684-881-23	Sequence 23, App	358	8	6.8	25	2	US-09-929-955-38	Sequence 38, Appl
286	8	6.8	9	1	US-08-146-028-285	Sequence 285, App	359	8	6.8	25	2	US-09-930-591-14	Sequence 14, Appl
287	8	6.8	9	1	US-08-146-028-291	Sequence 291, App	360	8	6.8	25	2	US-09-930-591-15	Sequence 15, Appl
288	8	6.8	9	1	US-08-146-028-292	Sequence 292, App	361	8	6.8	25	2	US-09-930-591-16	Sequence 16, Appl
289	8	6.8	9	1	US-08-146-028-298	Sequence 298, App	362	8	6.8	25	2	US-09-930-591-17	Sequence 17, Appl
290	8	6.8	9	2	US-08-723-425A-285	Sequence 285, App	363	8	6.8	25	2	US-09-930-591-18	Sequence 18, Appl
291	8	6.8	9	2	US-08-723-425A-291	Sequence 291, App	364	8	6.8	25	2	US-09-930-591-19	Sequence 19, Appl
292	8	6.8	9	2	US-08-723-425A-292	Sequence 292, App	365	8	6.8	25	2	US-09-930-591-20	Sequence 20, Appl
293	8	6.8	9	2	US-08-723-425A-298	Sequence 298, App	366	8	6.8	25	2	US-09-930-591-21	Sequence 21, Appl
294	8	6.8	9	2	US-09-112-206-285	Sequence 285, App	367	8	6.8	25	2	US-09-930-591-22	Sequence 22, Appl
295	8	6.8	9	2	US-09-112-206-291	Sequence 291, App	368	8	6.8	25	2	US-09-930-591-23	Sequence 23, Appl
296	8	6.8	9	2	US-09-112-206-292	Sequence 292, App	369	8	6.8	25	2	US-07-946-054-7	Sequence 7, Appli
297	8	6.8	9	2	US-09-112-206-298	Sequence 298, App	370	8	6.8	25	2	US-08-336-553A-17	Sequence 17, Appl
298	8	6.8	9	2	US-09-790-497A-275	Sequence 275, App	371	8	6.8	25	2	US-08-439-157-17	Sequence 17, Appl
299	8	6.8	9	2	US-09-576-824A-276	Sequence 276, App	372	8	6.8	25	2	US-09-439-157-17	Sequence 17, Appl
300	8	6.8	9	2	US-09-576-824A-275	Sequence 275, App	373	8	6.8	30	1	PCT-US93-08638-7	Sequence 7, Appli
301	8	6.8	9	2	US-09-680-497-285	Sequence 285, App	374	8	6.8	30	1	US-07-666-719-5	Sequence 5, Appli
302	8	6.8	9	2	US-09-680-497-291	Sequence 291, App	375	8	6.8	30	2	US-08-444-818-771	Sequence 771, App
303	8	6.8	9	2	US-09-680-497-292	Sequence 292, App	376	8	6.8	30	2	US-08-905-054B-13	Sequence 13, Appl
304	8	6.8	9	2	US-09-680-497-298	Sequence 298, App	377	8	6.8	30	4	PCT-US92-07813-13	Sequence 13, Appl
305	8	6.8	9	2	US-08-680-437-298	Sequence 298, App	378	8	6.8	31	1	US-07-946-054-6	Sequence 6, Appli
306	8	6.8	10	2	US-08-802-981-210	Sequence 210, App	379	8	6.8	30	1	US-08-083-947-22	Sequence 22, Appl
307	8	6.8	19	2	US-08-802-981-23	Sequence 23, App	380	8	6.8	41	2	US-08-530-530-2	Sequence 2, Appli
308	8	6.8	136	2	US-09-747-287A-136	Sequence 136, App	381	8	6.8	42	2	US-08-262-037-21	Sequence 21, Appl
309	8	6.8	19	2	US-09-394-019C-89	Sequence 89, Appl	382	8	6.8	47	1	PCT-US93-08638-6	Sequence 6, Appli
310	8	6.8	19	2	US-09-394-019C-334	Sequence 334, App	383	8	6.8	47	1	PCT-US94-07088-22	Sequence 22, Appl
311	8	6.8	20	1	US-08-466-975A-10	Sequence 10, Appl	384	8	6.8	47	1	PCT-US95-13660-2	Sequence 2, Appli
312	8	6.8	20	1	US-08-466-975A-11	Sequence 11, Appl	385	8	6.8	47	1	US-07-946-054-11	Sequence 11, Appl
313	8	6.8	20	1	US-08-391-671A-10	Sequence 10, Appl	386	8	6.8	47	4	PCT-US93-08638-11	Sequence 11, Appl
314	8	6.8	20	1	US-08-391-671A-11	Sequence 11, Appl	387	8	6.8	47	4	US-08-444-818-2	Sequence 2, Appli
315	8	6.8	20	1	US-08-853-623D-26	Sequence 26, Appl	388	8	6.8	47	4	US-08-905-054B-13	Sequence 13, Appl
316	8	6.8	20	1	US-08-853-623D-27	Sequence 27, Appl	389	8	6.8	50	1	US-07-946-054-11	Sequence 11, Appl
317	8	6.8	20	2	US-08-467-902A-10	Sequence 10, Appl	390	8	6.8	50	4	PCT-US93-08638-11	Sequence 11, Appl
318	8	6.8	20	2	US-08-467-902A-11	Sequence 11, Appl	391	8	6.8	51	2	US-08-905-054B-11	Sequence 11, Appl
319	8	6.8	20	2	US-09-208-966-47	Sequence 47, Appl	392	8	6.8	67	2		

393	8	6.8	67	4	PCT-US92-07813-11	Sequence 11, Appl	466	7	5.9	22	1	US-08-146-028-149	Sequence 149, App
394	8	6.8	117	2	US-08-444-818-10	Sequence 10, Appl	467	7	5.9	22	1	US-08-845-926-11	Sequence 11, Appl
395	8	6.8	117	2	US-08-905-054B-12	Sequence 12, Appl	468	7	5.9	22	1	US-08-845-926-12	Sequence 12, Appl
396	8	6.8	117	4	PCT-US92-07813-12	Sequence 12, Appl	469	7	5.9	22	2	US-08-723-425A-92	Sequence 92, Appl
397	8	6.8	128	2	US-08-444-818-8	Sequence 8, Appl	470	7	5.9	22	2	US-08-723-425A-93	Sequence 93, Appl
398	8	6.8	465	2	US-08-833-678A-2	Sequence 2, Appl	471	7	5.9	22	2	US-08-723-425A-149	Sequence 149, App
399	8	6.8	465	2	US-08-529-169A-2	Sequence 2, Appl	472	7	5.9	22	2	US-09-112-206-92	Sequence 92, Appl
400	8	6.8	465	2	US-09-483-799-2	Sequence 2, Appl	473	7	5.9	22	2	US-09-112-206-93	Sequence 93, Appl
401	8	6.8	512	2	US-08-867-611-58	Sequence 58, Appl	474	7	5.9	22	2	US-09-112-206-149	Sequence 149, App
402	8	6.8	512	2	US-09-690-359-58	Sequence 58, Appl	475	7	5.9	22	2	US-09-351-296-11	Sequence 11, Appl
403	8	6.8	613	2	US-10-104-966-6	Sequence 6, Appl	476	7	5.9	22	2	US-09-351-296-12	Sequence 12, Appl
404	8	6.8	613	2	US-09-929-955-6	Sequence 6, Appl	477	7	5.9	22	2	US-09-576-824A-92	Sequence 92, Appl
405	8	6.8	631	1	US-08-833-678A-1	Sequence 1, Appl	478	7	5.9	22	2	US-09-576-824A-93	Sequence 93, Appl
406	8	6.8	631	2	US-09-128-314-2	Sequence 2, Appl	479	7	5.9	22	2	US-09-680-497-92	Sequence 92, Appl
407	8	6.8	631	2	US-08-529-169A-1	Sequence 1, Appl	480	7	5.9	22	2	US-09-680-497-93	Sequence 93, Appl
408	8	6.8	631	2	US-09-483-799-1	Sequence 1, Appl	481	7	5.9	22	2	US-09-680-497-149	Sequence 149, App
409	8	6.8	632	2	US-09-929-955-29	Sequence 29, Appl	482	7	5.9	24	2	US-09-020-846-44	Sequence 44, Appl
410	8	6.8	632	2	US-09-930-591-12	Sequence 12, Appl	483	7	5.9	25	1	US-08-934-741A-23	Sequence 23, Appl
411	8	6.8	638	2	US-09-288-391-25	Sequence 25, Appl	484	7	5.9	30	2	US-08-336-553A-19	Sequence 19, Appl
412	8	6.8	781	2	US-08-867-611-4	Sequence 4, Appl	485	7	5.9	30	2	US-08-439-157-19	Sequence 19, Appl
413	8	6.8	781	2	US-09-690-359-4	Sequence 4, Appl	486	7	5.9	30	2	US-09-437-895-19	Sequence 19, Appl
414	8	6.8	781	4	PCT-US92-06965A-9	Sequence 9, Appl	487	7	5.9	34	1	US-08-934-741A-3	Sequence 3, Appl
415	8	6.8	1786	2	US-08-444-818-54	Sequence 54, Appl	488	7	5.9	40	1	US-07-946-054-8	Sequence 8, Appl
416	7	5.9	9	1	US-08-146-028-284	Sequence 284, App	489	7	5.9	40	1	US-08-530-550-35	Sequence 35, Appl
417	7	5.9	9	1	US-08-146-028-290	Sequence 290, App	490	7	5.9	40	1	US-08-530-550-35	Sequence 35, Appl
418	7	5.9	9	1	US-08-146-028-293	Sequence 293, App	491	7	5.9	40	1	US-08-530-550-35	Sequence 35, Appl
419	7	5.9	9	1	US-08-146-028-299	Sequence 299, App	492	7	5.9	40	4	PCT-US93-08638-8	Sequence 8, Appl
420	7	5.9	9	1	US-08-802-981-212	Sequence 212, App	493	7	5.9	53	2	US-09-270-767-39363	Sequence 39363, A
421	7	5.9	9	2	US-08-723-425A-284	Sequence 284, App	494	7	5.9	53	2	US-09-270-767-54580	Sequence 54580, A
422	7	5.9	9	2	US-08-723-425A-290	Sequence 290, App	495	7	5.9	87	1	US-08-685-764-4	Sequence 4, Appl
423	7	5.9	9	2	US-08-723-425A-293	Sequence 293, App	496	7	5.9	100	2	US-09-248-796A-17488	Sequence 17488, A
424	7	5.9	9	2	US-08-723-425A-299	Sequence 299, App	497	7	5.9	114	2	US-09-107-532A-6600	Sequence 6600, App
425	7	5.9	9	2	US-09-112-206-284	Sequence 284, App	498	7	5.9	128	2	US-09-878-281A-60	Sequence 60, Appl
426	7	5.9	9	2	US-09-112-206-290	Sequence 290, App	499	7	5.9	130	2	US-09-252-991A-25536	Sequence 25536, A
427	7	5.9	9	2	US-09-112-206-293	Sequence 293, App	500	7	5.9	131	2	US-09-164-615-34	Sequence 34, Appl
428	7	5.9	9	2	US-09-112-206-299	Sequence 299, App	501	7	5.9	146	2	US-09-602-787A-448	Sequence 448, App
429	7	5.9	9	2	US-09-790-497A-274	Sequence 274, App	502	7	5.9	158	2	US-09-602-787A-448	Sequence 448, App
430	7	5.9	9	2	US-09-790-497A-277	Sequence 277, App	503	7	5.9	163	2	US-09-252-991A-29832	Sequence 29832, A
431	7	5.9	9	2	US-09-576-824A-274	Sequence 274, App	504	7	5.9	179	2	US-09-687-363-10	Sequence 10, Appl
432	7	5.9	9	2	US-09-576-824A-277	Sequence 277, App	505	7	5.9	183	2	US-09-902-540-11456	Sequence 11456, A
433	7	5.9	9	2	US-09-680-497-284	Sequence 284, App	506	7	5.9	188	2	US-09-902-540-10069	Sequence 10069, A
434	7	5.9	9	2	US-09-680-497-290	Sequence 290, App	507	7	5.9	201	2	US-09-687-363-12	Sequence 12, Appl
435	7	5.9	9	2	US-09-680-497-293	Sequence 293, App	508	7	5.9	209	2	US-09-583-110-3075	Sequence 3075, App
436	7	5.9	9	2	US-09-680-497-299	Sequence 299, App	509	7	5.9	222	2	US-09-602-787A-442	Sequence 442, App
437	7	5.9	10	1	US-08-617-929-20	Sequence 20, Appl	510	7	5.9	227	2	US-09-107-433-3655	Sequence 3655, App
438	7	5.9	12	1	US-08-617-929-22	Sequence 22, Appl	511	7	5.9	254	2	US-09-248-796A-17487	Sequence 17487, A
439	7	5.9	14	1	US-08-439-747A-14	Sequence 14, Appl	512	7	5.9	260	2	US-10-104-966-8	Sequence 8, Appl
440	7	5.9	14	1	US-08-440-409B-14	Sequence 14, Appl	513	7	5.9	260	2	US-09-929-955-8	Sequence 8, Appl
441	7	5.9	14	1	US-08-617-929-21	Sequence 21, Appl	514	7	5.9	263	2	US-09-134-001C-4206	Sequence 4206, App
442	7	5.9	18	1	US-08-244-116B-4	Sequence 4, Appl	515	7	5.9	269	2	US-09-100-557-1	Sequence 1, Appl
443	7	5.9	18	2	US-08-802-981-96	Sequence 96, Appl	516	7	5.9	277	1	US-08-403-852D-18	Sequence 18, Appl
444	7	5.9	18	2	US-08-537-802-28	Sequence 28, Appl	517	7	5.9	277	2	US-08-510-646B-19	Sequence 19, Appl
445	7	5.9	139	2	US-09-747-287A-139	Sequence 139, App	518	7	5.9	277	2	US-09-231-818-18	Sequence 18, Appl
446	7	5.9	18	2	US-09-394-019C-92	Sequence 92, Appl	519	7	5.9	277	2	US-09-635-359B-18	Sequence 18, Appl
447	7	5.9	18	2	US-09-394-019C-337	Sequence 337, App	520	7	5.9	306	2	US-09-712-363-186	Sequence 186, App
448	7	5.9	19	2	US-08-802-981-99	Sequence 99, Appl	521	7	5.9	319	2	US-09-806-536A-3	Sequence 3, Appl
449	7	5.9	19	2	US-08-604-365-11	Sequence 11, Appl	522	7	5.9	367	2	US-09-545-072A-2	Sequence 2, Appl
450	7	5.9	19	2	US-09-689-678-11	Sequence 11, Appl	523	7	5.9	396	2	US-09-020-846-69	Sequence 69, Appl
451	7	5.9	19	2	US-09-747-287A-142	Sequence 142, App	524	7	5.9	396	2	US-10-133-007-11	Sequence 11, Appl
452	7	5.9	19	2	US-09-394-019C-95	Sequence 95, Appl	525	7	5.9	409	2	US-09-252-991A-30461	Sequence 30461, A
453	7	5.9	19	2	US-08-617-929-2	Sequence 2, Appl	526	7	5.9	421	2	US-09-020-846-68	Sequence 68, Appl
454	7	5.9	20	1	US-08-617-929-2	Sequence 2, Appl	527	7	5.9	433	2	US-09-328-352-7646	Sequence 7646, App
455	7	5.9	20	2	US-08-802-981-102	Sequence 102, App	528	7	5.9	438	2	US-09-489-039A-8263	Sequence 8263, App
456	7	5.9	20	2	US-09-790-497A-92	Sequence 92, Appl	529	7	5.9	457	2	US-09-252-991A-22682	Sequence 22682, A
457	7	5.9	20	2	US-09-790-497A-93	Sequence 93, Appl	530	7	5.9	477	2	US-09-902-540-11909	Sequence 11909, A
458	7	5.9	20	2	US-08-790-437A-149	Sequence 149, App	531	7	5.9	526	2	US-10-082-894-2	Sequence 2, Appl
459	7	5.9	20	2	US-09-576-824A-149	Sequence 149, App	532	7	5.9	526	2	US-10-082-894-4	Sequence 4, Appl
460	7	5.9	20	2	US-09-878-281A-278	Sequence 278, App	533	7	5.9	808	2	US-09-252-991A-31967	Sequence 31967, A
461	7	5.9	20	2	US-09-747-287A-145	Sequence 145, App	534	7	5.9	913	2	US-09-252-991A-29362	Sequence 29362, A
462	7	5.9	20	2	US-09-394-019C-98	Sequence 98, Appl	535	7	5.9	1036	2	US-10-104-047-2812	Sequence 2812, App
463	7	5.9	20	2	US-09-394-019C-98	Sequence 343, App	536	7	5.9	1116	2	US-09-543-681A-4379	Sequence 4379, App
464	7	5.9	22	1	US-08-146-028-92	Sequence 92, Appl	537	7	5.9	1255	2	US-09-605-703B-2408	Sequence 2408, App
465	7	5.9	22	1	US-08-146-028-93	Sequence 93, Appl	538	7	5.9	2940	2	US-10-226-629A-13	Sequence 13, Appl

539	7	5.9	3033	1	US-07-925-695-5	Sequence 5, Appli	612	6	5.1	20	2	US-09-011-961-7	Sequence 7, Appli
540	6	5.1	8	2	US-08-444-818-458	Sequence 458, App	613	6	5.1	20	2	US-09-208-966-48	Sequence 48, Appl
541	6	5.1	9	1	US-08-146-028-283	Sequence 283, App	614	6	5.1	20	2	US-09-775-052A-48	Sequence 48, Appl
542	6	5.1	9	1	US-08-146-028-289	Sequence 289, App	615	6	5.1	20	2	US-09-790-497A-94	Sequence 94, Appl
543	6	5.1	9	1	US-08-146-028-294	Sequence 294, App	616	6	5.1	20	2	US-09-878-281A-103	Sequence 103, App
544	6	5.1	9	1	US-08-146-028-300	Sequence 300, App	617	6	5.1	21	1	US-08-853-623D-1	Sequence 1, Appli
545	6	5.1	9	2	US-08-723-425A-283	Sequence 283, App	618	6	5.1	21	1	US-08-853-623D-2	Sequence 2, Appli
546	6	5.1	9	2	US-08-723-425A-289	Sequence 289, App	619	6	5.1	21	1	US-08-853-623D-6	Sequence 6, Appli
547	6	5.1	9	2	US-08-723-425A-294	Sequence 294, App	620	6	5.1	22	1	US-08-146-028-94	Sequence 94, Appl
548	6	5.1	9	2	US-08-723-425A-300	Sequence 300, App	621	6	5.1	22	1	US-08-723-425A-94	Sequence 94, Appl
549	6	5.1	9	2	US-09-112-206-283	Sequence 283, App	622	6	5.1	22	2	US-09-112-206-94	Sequence 94, Appl
550	6	5.1	9	2	US-09-112-206-289	Sequence 289, App	623	6	5.1	22	2	US-09-576-824A-94	Sequence 94, Appl
551	6	5.1	9	2	US-09-112-206-294	Sequence 294, App	624	6	5.1	22	2	US-09-680-497-94	Sequence 94, Appl
552	6	5.1	9	2	US-09-112-206-300	Sequence 300, App	625	6	5.1	24	1	US-08-484-635-127	Sequence 127, App
553	6	5.1	9	2	US-09-311-784A-257	Sequence 257, App	626	6	5.1	24	1	US-08-484-631-127	Sequence 127, App
554	6	5.1	9	2	US-09-311-784A-450	Sequence 450, App	627	6	5.1	24	1	US-08-827-570-127	Sequence 127, App
555	6	5.1	9	2	US-09-790-497A-273	Sequence 273, App	628	6	5.1	25	2	US-09-929-955-40	Sequence 40, Appl
556	6	5.1	9	2	US-09-790-497A-278	Sequence 278, App	629	6	5.1	25	2	US-09-930-531-25	Sequence 25, Appl
557	6	5.1	9	2	US-09-576-824A-273	Sequence 278, App	630	6	5.1	26	2	US-09-348-578-1	Sequence 1, Appli
558	6	5.1	9	2	US-09-576-824A-278	Sequence 278, App	631	6	5.1	26	2	US-09-699-684-1	Sequence 1, Appli
559	6	5.1	9	2	US-09-680-497-283	Sequence 283, App	632	6	5.1	27	2	US-09-348-578-2	Sequence 2, Appli
560	6	5.1	9	2	US-09-680-497-289	Sequence 283, App	633	6	5.1	27	2	US-09-348-578-10	Sequence 10, Appl
561	6	5.1	9	2	US-09-680-497-294	Sequence 294, App	634	6	5.1	27	2	US-09-699-684-3	Sequence 3, Appli
562	6	5.1	9	2	US-09-680-497-300	Sequence 300, App	635	6	5.1	27	2	US-09-699-684-2	Sequence 2, Appli
563	6	5.1	10	1	US-08-214-650-33	Sequence 33, Appl	636	6	5.1	28	2	US-09-699-684-10	Sequence 10, Appl
564	6	5.1	10	1	US-08-617-929-13	Sequence 13, Appl	637	6	5.1	28	2	US-09-348-578-3	Sequence 3, Appli
565	6	5.1	10	1	US-08-617-929-33	Sequence 33, Appl	638	6	5.1	28	2	US-09-348-578-11	Sequence 11, Appl
566	6	5.1	10	1	US-08-617-929-38	Sequence 38, Appl	639	6	5.1	28	2	US-09-348-578-19	Sequence 19, Appl
567	6	5.1	10	2	US-08-802-981-211	Sequence 211, App	640	6	5.1	28	2	US-09-699-684-3	Sequence 3, Appli
568	6	5.1	10	2	US-09-011-961-19	Sequence 19, Appl	641	6	5.1	28	2	US-09-699-684-11	Sequence 11, Appl
569	6	5.1	10	2	US-09-011-961-35	Sequence 35, Appl	642	6	5.1	28	2	US-09-699-684-19	Sequence 19, Appl
570	6	5.1	10	2	US-09-311-784A-435	Sequence 435, App	643	6	5.1	29	2	US-08-336-553A-18	Sequence 18, Appl
571	6	5.1	11	2	US-09-011-961-13	Sequence 13, Appl	644	6	5.1	29	2	US-09-348-578-4	Sequence 4, Appli
572	6	5.1	11	2	US-09-576-824A-528	Sequence 528, App	645	6	5.1	29	2	US-09-348-578-12	Sequence 12, Appl
573	6	5.1	12	1	US-08-617-929-17	Sequence 17, App	646	6	5.1	29	2	US-09-348-578-20	Sequence 20, Appl
574	6	5.1	12	1	US-08-853-623D-11	Sequence 11, Appl	647	6	5.1	29	2	US-08-439-157-18	Sequence 18, Appl
575	6	5.1	12	2	US-09-011-961-12	Sequence 12, Appl	648	6	5.1	29	2	US-09-437-895-18	Sequence 18, Appl
576	6	5.1	12	2	US-09-403-752A-78	Sequence 78, App	649	6	5.1	29	2	US-09-699-684-4	Sequence 4, Appli
577	6	5.1	12	2	US-09-551-151A-78	Sequence 78, App	650	6	5.1	29	2	US-09-699-684-12	Sequence 12, Appl
578	6	5.1	12	2	US-09-929-955-28	Sequence 28, App	651	6	5.1	29	2	US-09-699-684-20	Sequence 20, Appl
579	6	5.1	12	2	US-09-930-591-15	Sequence 15, App	652	6	5.1	30	2	US-08-336-553A-20	Sequence 20, Appl
580	6	5.1	13	1	US-08-439-747A-15	Sequence 15, App	653	6	5.1	30	2	US-09-348-578-5	Sequence 5, Appli
581	6	5.1	13	1	US-08-440-409B-15	Sequence 15, App	654	6	5.1	30	2	US-09-348-578-13	Sequence 13, Appl
582	6	5.1	13	1	US-08-853-623D-21	Sequence 21, Appl	655	6	5.1	30	2	US-09-348-578-21	Sequence 21, Appl
583	6	5.1	13	1	US-08-853-623D-30	Sequence 30, Appl	656	6	5.1	30	2	US-08-439-157-20	Sequence 20, Appl
584	6	5.1	13	2	US-09-011-961-11	Sequence 11, Appl	657	6	5.1	30	2	US-09-437-895-20	Sequence 20, Appl
585	6	5.1	13	2	US-09-719-261B-9	Sequence 9, Appli	658	6	5.1	30	2	US-09-699-684-5	Sequence 5, Appli
586	6	5.1	14	1	US-08-232-453A-7	Sequence 7, Appli	659	6	5.1	30	2	US-09-699-684-13	Sequence 13, Appl
587	6	5.1	14	1	US-08-232-453A-17	Sequence 17, Appl	660	6	5.1	30	2	US-09-699-684-21	Sequence 21, Appl
588	6	5.1	14	1	US-08-232-453A-25	Sequence 25, App	661	6	5.1	31	2	US-09-348-578-6	Sequence 6, Appli
589	6	5.1	14	1	US-08-232-453A-26	Sequence 26, App	662	6	5.1	31	2	US-09-348-578-14	Sequence 14, Appl
590	6	5.1	14	1	US-08-617-929-16	Sequence 16, App	663	6	5.1	31	2	US-09-348-578-22	Sequence 22, Appl
591	6	5.1	14	1	US-08-617-929-36	Sequence 36, App	664	6	5.1	31	2	US-09-699-684-6	Sequence 6, Appli
592	6	5.1	14	1	US-08-617-929-41	Sequence 41, App	665	6	5.1	31	2	US-09-699-684-14	Sequence 14, Appl
593	6	5.1	14	2	US-09-011-961-10	Sequence 10, App	666	6	5.1	31	2	US-09-699-684-22	Sequence 22, Appl
594	6	5.1	15	2	US-09-011-961-9	Sequence 9, Appli	667	6	5.1	32	2	US-09-348-578-7	Sequence 7, Appli
595	6	5.1	15	2	US-09-009-953-27	Sequence 27, App	668	6	5.1	32	2	US-09-348-578-15	Sequence 15, Appl
596	6	5.1	15	2	US-09-009-953-55	Sequence 55, App	669	6	5.1	32	2	US-09-348-578-23	Sequence 23, Appl
597	6	5.1	15	2	US-09-009-953-70	Sequence 70, App	670	6	5.1	32	2	US-09-699-684-7	Sequence 7, Appli
598	6	5.1	15	2	US-09-009-953-81	Sequence 81, App	671	6	5.1	32	2	US-09-699-684-15	Sequence 15, Appl
599	6	5.1	15	2	US-09-009-953-255	Sequence 255, App	672	6	5.1	32	2	US-09-699-684-23	Sequence 23, Appl
600	6	5.1	16	2	US-09-311-784A-425	Sequence 425, App	673	6	5.1	33	1	US-08-571-643A-7	Sequence 7, Appli
601	6	5.1	16	2	US-09-011-961-8	Sequence 8, Appli	674	6	5.1	33	1	US-08-571-643A-8	Sequence 8, Appli
602	6	5.1	18	2	US-08-537-802-32	Sequence 32, App	675	6	5.1	33	1	US-08-439-747A-29	Sequence 29, Appl
603	6	5.1	18	2	US-08-537-802-34	Sequence 34, App	676	6	5.1	33	1	US-08-439-747A-30	Sequence 30, Appl
604	6	5.1	19	2	US-08-802-981-94	Sequence 94, App	677	6	5.1	33	1	US-08-440-409B-30	Sequence 30, Appl
605	6	5.1	19	2	US-09-747-287A-137	Sequence 137, App	678	6	5.1	33	1	US-08-440-409B-31	Sequence 31, Appl
606	6	5.1	19	2	US-09-394-019C-90	Sequence 90, App	679	6	5.1	33	2	US-09-348-578-8	Sequence 8, Appli
607	6	5.1	19	2	US-09-394-019C-335	Sequence 335, App	680	6	5.1	33	2	US-09-348-578-16	Sequence 16, Appl
608	6	5.1	20	1	US-08-617-929-1	Sequence 1, Appli	681	6	5.1	33	2	US-09-348-578-24	Sequence 24, Appl
609	6	5.1	20	1	US-08-617-929-5	Sequence 5, Appli	682	6	5.1	33	2	US-09-699-684-8	Sequence 8, Appli
610	6	5.1	20	1	US-08-617-929-6	Sequence 6, Appli	683	6	5.1	33	2	US-09-699-684-16	Sequence 16, Appl
611	6	5.1	20	1	US-08-853-623D-8	Sequence 8, Appli	684	6	5.1	34	1	US-08-700-356-3	Sequence 3, Appli

685	6	5.1	34	1	US-08-936-865-3	Sequence 3, Appli	758	6	5.1	161	2	US-09-673-395A-156	Sequence 156, App
686	6	5.1	34	2	US-09-348-578-9	Sequence 9, Appli	759	6	5.1	162	1	US-08-319-704-6	Sequence 6, Appli
687	6	5.1	34	2	US-09-348-578-17	Sequence 17, Appl	760	6	5.1	162	2	US-09-605-703B-1008	Sequence 1008, Ap
688	6	5.1	34	2	US-09-348-578-25	Sequence 25, Appl	761	6	5.1	162	2	US-09-605-703B-1010	Sequence 1010, Ap
689	6	5.1	34	2	US-09-699-684-9	Sequence 9, Appli	762	6	5.1	166	2	US-09-252-991A-29685	Sequence 29685, A
690	6	5.1	34	2	US-09-699-684-17	Sequence 17, Appl	763	6	5.1	170	2	US-09-583-110-4668	Sequence 4668, Ap
691	6	5.1	34	2	US-09-699-684-25	Sequence 25, Appl	764	6	5.1	172	2	US-09-071-035-442	Sequence 442, App
692	6	5.1	35	2	US-09-348-578-18	Sequence 18, Appl	765	6	5.1	172	2	US-09-252-991A-26790	Sequence 26790, A
693	6	5.1	35	2	US-09-348-578-26	Sequence 26, Appl	766	6	5.1	172	2	US-10-206-576-442	Sequence 442, App
694	6	5.1	35	2	US-09-699-684-18	Sequence 18, Appl	767	6	5.1	173	2	US-08-956-171B-5218	Sequence 5218, Ap
695	6	5.1	35	2	US-09-699-684-26	Sequence 26, Appl	768	6	5.1	173	2	US-08-781-986A-5218	Sequence 5218, Ap
696	6	5.1	36	2	US-09-348-578-27	Sequence 27, Appl	769	6	5.1	176	2	US-09-902-540-14007	Sequence 14007, A
697	6	5.1	36	2	US-09-699-684-27	Sequence 27, Appl	770	6	5.1	186	2	US-09-252-991A-30079	Sequence 30079, A
698	6	5.1	36	2	US-09-585-682B-2	Sequence 2, Appli	771	6	5.1	186	2	US-09-134-000C-4177	Sequence 4177, Ap
699	6	5.1	37	2	US-09-902-540-16625	Sequence 16625, A	772	6	5.1	186	2	US-09-583-110-2827	Sequence 2827, Ap
700	6	5.1	38	1	US-08-118-270-185	Sequence 185, App	773	6	5.1	186	2	US-09-769-787-111	Sequence 111, App
701	6	5.1	38	4	PCT-US93-08528-185	Sequence 185, App	774	6	5.1	187	2	US-09-270-767-59012	Sequence 59012, A
702	6	5.1	40	1	US-08-617-929-9	Sequence 9, Appli	775	6	5.1	188	2	US-09-489-039A-13345	Sequence 13345, A
703	6	5.1	42	1	US-08-537-811-44	Sequence 44, Appl	776	6	5.1	190	2	US-09-328-352-4549	Sequence 4549, Ap
704	6	5.1	47	1	US-07-946-054-10	Sequence 10, Appl	777	6	5.1	193	2	US-09-134-000C-4429	Sequence 4429, Ap
705	6	5.1	47	4	PCT-US93-08638-10	Sequence 10, Appl	778	6	5.1	198	2	US-09-036-987A-23	Sequence 23, Appl
706	6	5.1	54	2	US-09-595-682B-4	Sequence 4, Appli	779	6	5.1	198	2	US-09-370-700-23	Sequence 23, Appl
707	6	5.1	57	2	US-09-905-223-407	Sequence 407, App	780	6	5.1	198	2	US-09-603-207-23	Sequence 23, Appl
708	6	5.1	63	1	US-08-685-764-1	Sequence 1, Appli	781	6	5.1	201	2	US-09-015-734-12	Sequence 12, Appl
709	6	5.1	63	1	US-08-685-764-3	Sequence 3, Appli	782	6	5.1	201	2	US-09-515-311-12	Sequence 12, Appl
710	6	5.1	70	2	US-09-513-999C-6688	Sequence 6688, Ap	783	6	5.1	201	2	US-10-434-817-12	Sequence 12, Appl
711	6	5.1	71	2	US-09-248-796A-25370	Sequence 25370, A	784	6	5.1	202	2	US-09-540-236-2414	Sequence 2414, Ap
712	6	5.1	72	2	US-09-513-999C-7755	Sequence 7755, Ap	785	6	5.1	202	2	US-09-107-433-3144	Sequence 3144, Ap
713	6	5.1	78	2	US-09-902-540-13136	Sequence 13136, A	786	6	5.1	203	2	US-09-252-991A-24035	Sequence 24035, A
714	6	5.1	87	2	US-08-685-764-2	Sequence 2, Appli	787	6	5.1	214	2	US-09-257-583-15	Sequence 15, Appl
715	6	5.1	91	2	US-08-479-078-1	Sequence 1, Appli	788	6	5.1	218	2	US-09-252-991A-19931	Sequence 19931, A
716	6	5.1	97	2	US-09-107-532A-3655	Sequence 3655, Ap	789	6	5.1	221	2	US-09-482-273-119	Sequence 119, App
717	6	5.1	98	1	US-08-308-086-3	Sequence 3, Appli	790	6	5.1	221	2	US-09-583-110-5016	Sequence 5016, Ap
718	6	5.1	98	2	US-08-975-040-21	Sequence 21, Appl	791	6	5.1	221	2	US-09-769-787-86	Sequence 86, Appl
719	6	5.1	99	1	US-08-167-035-19	Sequence 19, Appl	792	6	5.1	225	1	US-08-886-765-2	Sequence 2, Appli
720	6	5.1	99	1	US-08-208-887A-19	Sequence 19, Appl	793	6	5.1	225	2	US-09-115-660-2	Sequence 2, Appli
721	6	5.1	99	1	US-08-208-887A-50	Sequence 50, Appl	794	6	5.1	226	2	US-09-252-991A-17728	Sequence 17728, A
722	6	5.1	99	1	US-08-539-005-19	Sequence 19, Appl	795	6	5.1	227	2	US-09-071-035-476	Sequence 476, App
723	6	5.1	99	2	US-09-280-598-21	Sequence 21, Appl	796	6	5.1	227	2	US-10-206-576-476	Sequence 476, App
724	6	5.1	100	2	US-09-087-465-37	Sequence 37, Appl	797	6	5.1	227	2	US-09-605-703B-818	Sequence 818, App
725	6	5.1	101	2	US-09-513-999C-6858	Sequence 6858, Ap	798	6	5.1	229	2	US-09-252-991A-30629	Sequence 30629, A
726	6	5.1	105	1	US-08-820-754-22	Sequence 22, Appl	799	6	5.1	230	2	US-09-107-532A-7215	Sequence 7215, Ap
727	6	5.1	105	2	US-08-956-652-22	Sequence 22, Appl	800	6	5.1	231	2	US-09-328-352-7928	Sequence 7928, Ap
728	6	5.1	105	2	US-08-956-869-22	Sequence 22, Appl	801	6	5.1	231	2	US-09-949-016-7831	Sequence 7831, Ap
729	6	5.1	105	2	US-09-948-547-22	Sequence 22, Appl	802	6	5.1	232	2	US-09-270-767-42502	Sequence 42502, A
730	6	5.1	105	2	US-09-732-210-371	Sequence 371, App	803	6	5.1	233	2	US-09-107-433-5113	Sequence 5113, Ap
731	6	5.1	105	2	US-09-732-210-1056	Sequence 1056, Ap	804	6	5.1	236	2	US-09-015-734-7	Sequence 7, Appli
732	6	5.1	105	2	US-08-212-185-22	Sequence 22, Appl	805	6	5.1	236	2	US-09-515-311-7	Sequence 7, Appli
733	6	5.1	110	2	US-08-858-207A-273	Sequence 273, App	806	6	5.1	236	2	US-10-434-817-7	Sequence 7, Appli
734	6	5.1	110	2	US-09-417-251A-2	Sequence 2, Appli	807	6	5.1	237	2	US-09-248-796A-18049	Sequence 18049, A
735	6	5.1	110	2	US-09-417-251A-2	Sequence 2, Appli	808	6	5.1	238	2	US-09-252-991A-23387	Sequence 23387, A
736	6	5.1	112	2	US-09-504-358-24	Sequence 24, Appl	809	6	5.1	238	2	US-09-902-540-12387	Sequence 12387, A
737	6	5.1	112	2	US-09-954-314-24	Sequence 24, Appl	810	6	5.1	240	2	US-09-252-991A-28112	Sequence 28112, A
738	6	5.1	112	2	US-10-230-562-24	Sequence 24, Appl	811	6	5.1	243	2	US-09-252-991A-19639	Sequence 19639, A
739	6	5.1	115	2	US-09-583-110-3398	Sequence 3398, Ap	812	6	5.1	246	2	US-09-252-991A-27611	Sequence 27611, A
740	6	5.1	120	2	US-09-107-433-3440	Sequence 3440, Ap	813	6	5.1	253	2	US-09-252-991A-24362	Sequence 24362, A
741	6	5.1	123	2	US-09-248-796A-17484	Sequence 17484, A	814	6	5.1	255	2	US-09-015-734-2	Sequence 2, Appli
742	6	5.1	127	2	US-08-811-481-5	Sequence 5, Appli	815	6	5.1	255	2	US-09-515-311-2	Sequence 2, Appli
743	6	5.1	127	2	US-08-876-527-5	Sequence 5, Appli	816	6	5.1	255	2	US-09-515-311-2	Sequence 2, Appli
744	6	5.1	127	2	US-09-902-540-14345	Sequence 14345, A	817	6	5.1	255	2	US-09-902-540-13380	Sequence 13380, A
745	6	5.1	129	2	US-09-134-000C-3900	Sequence 3900, Ap	818	6	5.1	258	2	US-10-434-817-2	Sequence 2, Appli
746	6	5.1	134	2	US-09-308-386A-4	Sequence 4, Appli	819	6	5.1	261	2	US-09-252-991A-20260	Sequence 20260, A
747	6	5.1	138	2	US-09-902-540-14756	Sequence 14756, A	820	6	5.1	262	2	US-09-252-991A-20609	Sequence 20609, A
748	6	5.1	143	2	US-09-328-352-4915	Sequence 4915, Ap	821	6	5.1	267	2	US-09-252-991A-32324	Sequence 32324, A
749	6	5.1	144	2	US-09-328-352-7270	Sequence 7270, Ap	822	6	5.1	267	2	US-08-311-731A-114	Sequence 114, App
750	6	5.1	146	2	US-09-902-540-12071	Sequence 12071, A	823	6	5.1	269	2	US-09-813-453B-15	Sequence 15, Appl
751	6	5.1	151	2	US-09-252-991A-32059	Sequence 32059, A	824	6	5.1	271	2	US-09-270-767-32500	Sequence 32500, A
752	6	5.1	151	2	US-08-270-767-49120	Sequence 49120, A	825	6	5.1	271	2	US-09-252-991A-18538	Sequence 18538, A
753	6	5.1	155	2	US-08-336-553A-2	Sequence 2, Appli	826	6	5.1	280	2	US-09-489-039A-8452	Sequence 8452, Ap
754	6	5.1	155	2	US-08-439-157-2	Sequence 2, Appli	827	6	5.1	282	2	US-09-583-110-5207	Sequence 5207, Ap
755	6	5.1	155	2	US-09-437-895-2	Sequence 2, Appli	828	6	5.1	283	2	US-09-489-039A-9220	Sequence 9220, Ap
756	6	5.1	157	2	US-09-252-991A-24143	Sequence 24143, A	829	6	5.1	285	2	US-09-252-991A-20552	Sequence 20552, A
757	6	5.1	161	2	US-09-252-991A-23802	Sequence 23802, A	830	6	5.1	287	2	US-09-328-352-7096	Sequence 7096, Ap

831	6	5.1	287	2	US-09-540-236-2477	Sequence 2477, Ap	904	399	2	US-09-330-235-20	Sequence 20, Appl
832	6	5.1	287	2	US-09-502-540-13218	Sequence 13218, A	905	402	1	US-08-403-852D-19	Sequence 19, Appl
833	6	5.1	288	2	US-09-252-991A-27979	Sequence 27979, A	906	402	2	US-08-510-646B-20	Sequence 20, Appl
834	6	5.1	288	2	US-09-949-016-11675	Sequence 11675, A	907	402	2	US-09-231-818-19	Sequence 19, Appl
835	6	5.1	289	2	US-09-071-035-480	Sequence 480, App	908	402	2	US-09-635-359B-19	Sequence 19, Appl
836	6	5.1	289	2	US-10-206-576-480	Sequence 480, App	909	403	2	US-09-438-185A-567	Sequence 567, App
837	6	5.1	290	2	US-09-328-352-7390	Sequence 7390, App	910	407	2	US-09-328-352-5318	Sequence 5318, App
838	6	5.1	291	2	US-09-252-991A-22826	Sequence 22826, A	911	408	2	US-09-712-363-268	Sequence 268, App
839	6	5.1	296	2	US-09-724-623-103	Sequence 103, App	912	413	2	US-09-902-540-13562	Sequence 13562, A
840	6	5.1	298	2	US-08-811-481-4	Sequence 4, Appli	913	415	2	US-09-489-039A-10457	Sequence 10457, A
841	6	5.1	298	2	US-09-876-527-4	Sequence 4, Appli	914	415	2	US-09-740-288A-28	Sequence 28, Appl
842	6	5.1	299	2	US-09-252-991A-22835	Sequence 22835, A	915	417	2	US-08-462-467B-6	Sequence 6, Appli
843	6	5.1	302	2	US-09-252-991A-26364	Sequence 26364, A	916	418	2	US-09-543-681A-7435	Sequence 7435, Ap
844	6	5.1	304	2	US-09-252-991A-24026	Sequence 24026, A	917	418	2	US-09-614-912-196	Sequence 196, App
845	6	5.1	304	2	US-09-328-352-5720	Sequence 5720, Ap	918	427	2	US-08-065-844A-2	Sequence 2, Appli
846	6	5.1	304	2	US-09-328-352-5720	Sequence 5720, Ap	919	427	2	US-10-002-278-2	Sequence 2, Appli
847	6	5.1	304	2	US-09-248-796A-20160	Sequence 20160, A	920	429	2	US-09-252-991A-16841	Sequence 16841, A
848	6	5.1	306	2	US-09-160-975A-2	Sequence 2, Appli	921	430	2	US-09-134-000C-4819	Sequence 4819, Ap
849	6	5.1	306	2	US-09-252-991A-26798	Sequence 26798, A	922	432	2	US-09-489-039A-13901	Sequence 13901, A
850	6	5.1	316	2	US-09-248-796A-20451	Sequence 20451, A	923	432	2	US-09-902-540-10023	Sequence 10023, A
851	6	5.1	319	2	US-09-252-991A-23561	Sequence 23561, A	924	432	2	US-09-902-540-15774	Sequence 15774, A
852	6	5.1	321	2	US-09-583-110-3120	Sequence 3120, Ap	925	432	2	US-09-605-703B-1244	Sequence 1244, A
853	6	5.1	321	2	US-09-583-110-3120	Sequence 3120, Ap	926	433	2	US-09-073-030-220	Sequence 220, App
854	6	5.1	322	2	US-09-107-433-3055	Sequence 3055, Ap	927	434	2	US-09-252-991A-22522	Sequence 22522, A
855	6	5.1	323	2	US-09-902-540-13351	Sequence 13351, A	928	437	2	US-09-252-991A-24390	Sequence 24390, A
856	6	5.1	323	2	US-10-272-490-42	Sequence 42, Appl	929	439	2	US-09-902-540-15543	Sequence 15543, A
857	6	5.1	326	2	US-09-252-991A-27058	Sequence 27058, A	930	440	2	US-10-104-047-3856	Sequence 3856, Ap
858	6	5.1	327	2	US-09-248-796A-24407	Sequence 24407, A	931	441	2	US-09-252-991A-17727	Sequence 17727, A
859	6	5.1	330	2	US-09-724-797-28	Sequence 28, Appl	932	444	2	US-08-822-774-52	Sequence 52, Appl
860	6	5.1	331	2	US-09-252-991A-26165	Sequence 26165, A	933	444	2	US-09-632-711-52	Sequence 52, Appl
861	6	5.1	332	2	US-09-902-540-9720	Sequence 9720, Ap	934	444	2	US-09-632-703B-52	Sequence 52, Appl
862	6	5.1	333	2	US-09-758-759-204	Sequence 204, App	935	444	2	US-09-632-702-52	Sequence 52, Appl
863	6	5.1	335	2	US-09-570-856B-15	Sequence 15, Appl	936	444	2	US-09-252-991A-20496	Sequence 20496, A
864	6	5.1	337	2	US-09-489-039A-10149	Sequence 10149, A	937	444	2	US-09-393-003-52	Sequence 52, Appl
865	6	5.1	340	2	US-09-252-991A-24373	Sequence 24373, A	938	445	1	US-09-047-026A-6	Sequence 6, Appli
866	6	5.1	342	2	US-09-252-991A-27223	Sequence 27223, A	939	446	2	US-09-543-681A-5700	Sequence 5700, Ap
867	6	5.1	342	2	US-09-489-039A-8662	Sequence 8662, Ap	940	446	2	US-09-902-540-12215	Sequence 12215, A
868	6	5.1	346	2	US-09-543-681A-4974	Sequence 4974, Ap	941	450	2	US-09-071-035-356	Sequence 356, App
869	6	5.1	346	2	US-09-248-796A-16397	Sequence 16397, A	942	450	2	US-10-206-576-256	Sequence 256, App
870	6	5.1	347	2	US-09-489-039A-9489	Sequence 9489, Ap	943	452	2	US-09-248-796A-20296	Sequence 20296, A
871	6	5.1	349	2	US-09-252-991A-20858	Sequence 20858, A	944	453	2	US-09-949-016-10146	Sequence 10146, A
872	6	5.1	350	2	US-09-489-039A-10384	Sequence 10384, A	945	454	2	US-09-252-991A-24836	Sequence 24836, A
873	6	5.1	354	2	US-09-489-039A-13104	Sequence 13104, A	946	455	2	US-09-107-532A-6473	Sequence 6473, Ap
874	6	5.1	359	2	US-09-602-787A-466	Sequence 466, App	947	459	2	US-09-252-991A-22668	Sequence 22668, A
875	6	5.1	360	2	US-09-542-767A-1	Sequence 1, Appli	948	459	2	US-09-949-016-7329	Sequence 7329, Ap
876	6	5.1	366	2	US-09-252-991A-29984	Sequence 29984, A	949	462	2	US-09-607-248B-3	Sequence 3, Appli
877	6	5.1	367	2	US-09-198-452A-604	Sequence 604, App	950	465	2	US-09-949-016-10240	Sequence 10240, A
878	6	5.1	367	2	US-09-794-534-2	Sequence 2, Appli	951	465	2	US-09-902-540-10552	Sequence 10552, A
879	6	5.1	368	2	US-09-252-991A-20724	Sequence 20724, A	952	466	2	US-09-252-991A-30829	Sequence 30829, A
880	6	5.1	371	2	US-09-252-991A-20081	Sequence 20081, A	953	469	2	US-09-543-681A-7068	Sequence 7068, Ap
881	6	5.1	371	2	US-09-826-509-457	Sequence 457, App	954	470	2	US-10-033-109-14	Sequence 14, Appl
882	6	5.1	372	2	US-09-252-991A-31788	Sequence 31788, A	955	472	2	US-09-489-039A-13479	Sequence 13479, A
883	6	5.1	374	2	US-09-740-288A-12	Sequence 12, Appl	956	472	2	US-09-902-540-14925	Sequence 14925, A
884	6	5.1	376	1	US-08-883-534-5	Sequence 5, Appli	957	473	2	US-09-248-796A-16084	Sequence 16084, A
885	6	5.1	376	2	US-09-204-764-5	Sequence 5, Appli	958	479	2	US-09-252-991A-19246	Sequence 19246, A
886	6	5.1	376	2	US-08-811-481-7	Sequence 7, Appli	959	483	2	US-09-270-767-45923	Sequence 45923, A
887	6	5.1	376	2	US-09-876-527-7	Sequence 7, Appli	960	483	2	US-09-902-540-9804	Sequence 9804, Ap
888	6	5.1	379	2	US-09-107-532A-5519	Sequence 5519, Ap	961	487	2	US-09-270-767-57297	Sequence 57297, A
889	6	5.1	380	2	US-09-097-889-25	Sequence 25, Appl	962	488	2	US-09-489-039A-10133	Sequence 10133, A
890	6	5.1	380	2	US-09-098-079-25	Sequence 25, Appl	963	490	2	US-09-252-991A-26250	Sequence 26250, A
891	6	5.1	384	2	US-09-902-540-10126	Sequence 10126, A	964	492	2	US-09-543-681A-6264	Sequence 6264, Ap
892	6	5.1	385	2	US-09-270-767-43634	Sequence 43634, A	965	493	2	US-09-252-991A-23214	Sequence 23214, A
893	6	5.1	385	2	US-09-712-363-148	Sequence 148, App	966	496	2	US-09-807-258-2	Sequence 2, Appli
894	6	5.1	385	2	US-09-938-901A-62	Sequence 62, Appl	967	496	2	US-08-807-258-4	Sequence 4, Appli
895	6	5.1	389	2	US-09-328-352-6823	Sequence 6823, Ap	968	499	2	US-09-489-039A-10027	Sequence 10027, A
896	6	5.1	390	2	US-09-902-540-16621	Sequence 16621, A	969	500	2	US-09-328-352-8183	Sequence 8183, Ap
897	6	5.1	396	2	US-09-489-039A-8995	Sequence 8995, Ap	970	504	2	US-09-949-016-6538	Sequence 6538, Ap
898	6	5.1	397	2	US-09-489-039A-13498	Sequence 13498, A	971	505	2	US-09-949-016-6538	Sequence 6538, Ap
899	6	5.1	398	1	US-08-834-655-4	Sequence 10656, A	972	506	2	US-09-252-991A-19050	Sequence 19050, A
900	6	5.1	399	2	US-08-834-655-4	Sequence 4, Appli	973	510	2	US-09-252-991A-29467	Sequence 29467, App
901	6	5.1	399	2	US-08-834-655A-4	Sequence 4, Appli	974	510	2	US-09-712-363-227	Sequence 227, App
902	6	5.1	399	2	US-09-363-574-4	Sequence 4, Appli	975	512	2	US-09-902-540-13924	Sequence 13924, A
903	6	5.1	399	2	US-09-363-526-4	Sequence 4, Appli	976	513	2	US-09-252-991A-30773	Sequence 30773, A

977	6	5.1	513	2	US-09-833-745-56	Sequence 56, Appl	1050	6	5.1	657	2	US-09-949-016-6174	Sequence 6174, Ap
978	6	5.1	513	2	US-09-833-745-59	Sequence 59, Appl	1051	6	5.1	657	2	US-09-833-745-36	Sequence 36, Appl
979	6	5.1	513	2	US-09-833-745-60	Sequence 60, Appl	1052	6	5.1	657	2	US-09-833-745-37	Sequence 37, Appl
980	6	5.1	513	2	US-09-833-745-61	Sequence 61, Appl	1053	6	5.1	662	2	US-09-538-092-1325	Sequence 1325, Ap
981	6	5.1	515	2	US-09-252-991A-20224	Sequence 20224, A	1054	6	5.1	662	2	US-09-949-016-6619	Sequence 6619, Ap
982	6	5.1	520	2	US-09-252-991A-28903	Sequence 28903, A	1055	6	5.1	665	2	US-09-949-016-10710	Sequence 10710, A
983	6	5.1	521	2	US-09-949-016-8809	Sequence 8809, Ap	1056	6	5.1	666	2	US-10-104-047-2217	Sequence 2217, Ap
984	6	5.1	524	2	US-09-252-991A-18092	Sequence 18092, A	1057	6	5.1	671	2	US-09-252-991A-18450	Sequence 18450, A
985	6	5.1	524	2	US-09-242-913B-15	Sequence 15, Appl	1058	6	5.1	673	2	US-09-063-950-2	Sequence 2, Appl
986	6	5.1	529	2	US-09-252-991A-26641	Sequence 26641, A	1059	6	5.1	673	2	US-09-991-181-52	Sequence 52, Appl
987	6	5.1	530	2	US-08-975-762-73	Sequence 73, Appl	1060	6	5.1	673	2	US-09-990-444-52	Sequence 52, Appl
988	6	5.1	530	2	US-09-295-028-73	Sequence 73, Appl	1061	6	5.1	673	2	US-09-997-333-52	Sequence 52, Appl
989	6	5.1	530	2	US-09-106-582-73	Sequence 73, Appl	1062	6	5.1	673	2	US-09-992-598-52	Sequence 52, Appl
990	6	5.1	530	2	US-09-159-469-73	Sequence 73, Appl	1063	6	5.1	675	2	US-09-248-796A-20699	Sequence 20699, A
991	6	5.1	530	2	US-09-693-542-73	Sequence 73, Appl	1064	6	5.1	676	2	US-09-252-991A-32706	Sequence 32706, A
992	6	5.1	530	2	US-09-444-711A-4	Sequence 4, Appl	1065	6	5.1	678	2	US-09-949-016-10001	Sequence 10001, A
993	6	5.1	532	1	US-08-339-152A-32	Sequence 32, Appl	1066	6	5.1	690	2	US-09-328-352-5182	Sequence 5182, Ap
994	6	5.1	533	1	US-07-820-011A-2	Sequence 2, Appl	1067	6	5.1	693	2	US-09-252-991A-24059	Sequence 24059, A
995	6	5.1	533	2	US-09-470-881-3	Sequence 3, Appl	1068	6	5.1	707	2	US-09-252-991A-24045	Sequence 24045, A
996	6	5.1	533	4	PCT-US93-00445-2	Sequence 2, Appl	1069	6	5.1	715	2	US-09-902-540-9752	Sequence 9752, Ap
997	6	5.1	534	2	US-09-902-540-13260	Sequence 13260, A	1070	6	5.1	724	2	US-09-252-991A-21494	Sequence 21494, A
998	6	5.1	536	1	US-07-820-011A-4	Sequence 4, Appl	1071	6	5.1	728	2	US-09-252-991A-28459	Sequence 28459, A
999	6	5.1	536	2	US-08-436-509A-13	Sequence 13, Appl	1072	6	5.1	731	2	US-09-134-000C-4355	Sequence 4355, Ap
1000	6	5.1	536	2	US-08-232-545-13	Sequence 13, Appl	1073	6	5.1	736	2	US-09-252-991A-19048	Sequence 19048, A
1001	6	5.1	536	2	US-09-444-711A-2	Sequence 2, Appl	1074	6	5.1	748	1	US-08-997-080-154	Sequence 154, App
1002	6	5.1	536	2	US-09-929-266-10	Sequence 10, Appl	1075	6	5.1	748	1	US-08-997-362-154	Sequence 154, App
1003	6	5.1	536	2	US-09-977-261-13	Sequence 13, Appl	1076	6	5.1	748	2	US-09-095-855-154	Sequence 154, App
1004	6	5.1	536	4	PCT-US93-00445-4	Sequence 4, Appl	1077	6	5.1	748	2	US-09-324-542-154	Sequence 154, App
1005	6	5.1	536	4	PCT-US95-05008-13	Sequence 13, Appl	1078	6	5.1	748	2	US-09-205-426-154	Sequence 154, App
1006	6	5.1	537	2	US-09-489-039A-10180	Sequence 10180, A	1079	6	5.1	752	2	US-10-104-047-1975	Sequence 1975, Ap
1007	6	5.1	537	2	US-09-949-016-10282	Sequence 10282, A	1080	6	5.1	761	2	US-09-585-858-27	Sequence 27, Appl
1008	6	5.1	539	2	US-09-264-737-1	Sequence 1, Appl	1081	6	5.1	761	2	US-10-270-878-27	Sequence 27, Appl
1009	6	5.1	540	2	US-09-252-991A-26410	Sequence 26410, A	1082	6	5.1	764	1	US-08-424-567-2	Sequence 2, Appl
1010	6	5.1	542	2	US-09-902-540-12859	Sequence 12859, A	1083	6	5.1	764	1	US-08-711-928-2	Sequence 2, Appl
1011	6	5.1	545	2	US-09-252-991A-31086	Sequence 31086, A	1084	6	5.1	764	2	US-09-184-937-2	Sequence 2, Appl
1012	6	5.1	547	2	US-09-066-047-2	Sequence 2, Appl	1085	6	5.1	773	2	US-09-949-016-10285	Sequence 10285, A
1013	6	5.1	547	2	US-09-252-991A-17363	Sequence 17363, A	1086	6	5.1	780	2	US-09-949-016-6372	Sequence 6372, Ap
1014	6	5.1	555	2	US-09-252-991A-27459	Sequence 27459, A	1087	6	5.1	781	2	US-09-252-991A-23353	Sequence 23353, A
1015	6	5.1	556	2	US-08-462-467B-24	Sequence 24, Appl	1088	6	5.1	781	2	US-10-104-047-2979	Sequence 2979, Ap
1016	6	5.1	556	2	US-08-462-467B-26	Sequence 26, Appl	1089	6	5.1	790	2	US-09-949-016-11220	Sequence 11220, A
1017	6	5.1	558	2	US-09-252-991A-24047	Sequence 24047, A	1090	6	5.1	798	2	US-09-252-991A-30781	Sequence 30781, A
1018	6	5.1	561	2	US-09-252-991A-16726	Sequence 16726, A	1091	6	5.1	803	2	US-09-252-991A-23614	Sequence 23614, A
1019	6	5.1	564	2	US-09-252-991A-22615	Sequence 22615, A	1092	6	5.1	814	2	US-09-252-991A-22323	Sequence 22323, A
1020	6	5.1	565	2	US-09-489-039A-8903	Sequence 8903, Ap	1093	6	5.1	815	2	US-09-902-540-10502	Sequence 10502, A
1021	6	5.1	566	2	US-09-264-737-2	Sequence 2, Appl	1094	6	5.1	818	2	US-08-811-481-22	Sequence 22, Appl
1022	6	5.1	579	2	US-09-902-540-12155	Sequence 12155, A	1095	6	5.1	818	2	US-09-876-527-22	Sequence 22, Appl
1023	6	5.1	579	2	US-09-252-991A-26166	Sequence 26166, A	1096	6	5.1	832	2	US-09-252-991A-17367	Sequence 17367, A
1024	6	5.1	582	1	US-08-261-086-8	Sequence 8, Appl	1097	6	5.1	844	2	US-09-931-258-19	Sequence 19, Appl
1025	6	5.1	586	2	US-09-252-991A-32972	Sequence 32972, A	1098	6	5.1	851	2	US-09-252-991A-24773	Sequence 24773, A
1026	6	5.1	587	2	US-09-252-991A-31749	Sequence 31749, A	1099	6	5.1	861	2	US-09-826-312A-10	Sequence 10, Appl
1027	6	5.1	590	2	US-08-975-762-64	Sequence 64, Appl	1100	6	5.1	861	2	US-09-542-497A-10	Sequence 10, Appl
1028	6	5.1	590	2	US-09-295-028-64	Sequence 64, Appl	1101	6	5.1	861	2	US-10-108-767-10	Sequence 10, Appl
1029	6	5.1	590	2	US-09-106-582-64	Sequence 64, Appl	1102	6	5.1	866	2	US-09-252-991A-32568	Sequence 32568, A
1030	6	5.1	590	2	US-09-159-469-64	Sequence 64, Appl	1103	6	5.1	873	1	US-09-071-035-246	Sequence 246, App
1031	6	5.1	590	2	US-09-693-542-64	Sequence 64, Appl	1104	6	5.1	873	1	US-08-571-758-6	Sequence 6, Appl
1032	6	5.1	590	2	US-09-902-540-9740	Sequence 9740, Ap	1105	6	5.1	873	1	US-08-909-983-6	Sequence 6, Appl
1033	6	5.1	598	2	US-09-866-028-69	Sequence 69, Appl	1106	6	5.1	878	2	US-09-902-540-10577	Sequence 10577, A
1034	6	5.1	598	2	US-09-944-457-69	Sequence 69, Appl	1107	6	5.1	894	2	US-09-071-035-248	Sequence 248, App
1035	6	5.1	598	2	US-09-945-584-69	Sequence 69, Appl	1108	6	5.1	894	2	US-10-206-576-248	Sequence 248, App
1036	6	5.1	598	2	US-09-944-944-69	Sequence 69, Appl	1109	6	5.1	921	2	US-09-252-991A-17504	Sequence 17504, A
1037	6	5.1	598	2	US-09-945-587-69	Sequence 69, Appl	1110	6	5.1	937	2	US-09-252-991A-19108	Sequence 19108, A
1038	6	5.1	603	2	US-08-482-677-8	Sequence 8, Appl	1111	6	5.1	956	2	US-09-252-991A-28002	Sequence 28002, A
1039	6	5.1	603	2	US-10-033-174-8	Sequence 8, Appl	1112	6	5.1	957	2	US-09-252-991A-32568	Sequence 32568, A
1040	6	5.1	604	1	US-08-635-137-2	Sequence 2, Appl	1113	6	5.1	962	2	US-09-071-035-246	Sequence 246, App
1041	6	5.1	613	2	US-09-136-981-2	Sequence 2, Appl	1114	6	5.1	962	2	US-09-071-035-250	Sequence 250, App
1042	6	5.1	613	2	US-09-252-991A-25899	Sequence 25899, A	1115	6	5.1	962	2	US-09-071-035-254	Sequence 254, App
1043	6	5.1	621	2	US-09-252-991A-19281	Sequence 19281, A	1116	6	5.1	962	2	US-09-071-035-470	Sequence 470, App
1044	6	5.1	627	2	US-09-252-991A-32694	Sequence 32694, A	1117	6	5.1	962	2	US-09-071-035-474	Sequence 474, App
1045	6	5.1	628	2	US-09-252-991A-32414	Sequence 32414, A	1118	6	5.1	962	2	US-09-071-035-478	Sequence 478, App
1046	6	5.1	635	2	US-09-833-745-38	Sequence 38, Appl	1119	6	5.1	962	2	US-10-206-576-246	Sequence 246, App
1047	6	5.1	641	2	US-08-811-481-2	Sequence 2, Appl	1120	6	5.1	962	2	US-10-206-576-250	Sequence 250, App
1048	6	5.1	641	2	US-09-876-527-2	Sequence 2, Appl	1121	6	5.1	962	2	US-10-206-576-254	Sequence 254, App
1049	6	5.1	656	2	US-09-252-991A-18768	Sequence 18768, A	1122	6	5.1	962	2	US-10-206-576-470	Sequence 470, App

1123	6	5.1	962	2	US-10-206-576-474	Sequence 474, App	1196	5	4.2	5	2	US-09-104-337A-411	Sequence 411, App
1124	6	5.1	962	2	US-10-206-576-478	Sequence 478, App	1197	5	4.2	5	2	US-09-719-261B-6	Sequence 6, Appli
1125	6	5.1	969	1	US-08-548-159-1	Sequence 1, Appli	1198	5	4.2	5	2	US-09-719-261B-2	Sequence 2, Appli
1126	6	5.1	969	2	US-09-252-991A-36985	Sequence 26985, A	1199	5	4.2	5	2	US-09-719-261B-5	Sequence 5, Appli
1127	6	5.1	970	2	US-09-134-000C-5691	Sequence 5691, Ap	1200	5	4.2	5	2	US-08-604-365-8	Sequence 8, Appli
1128	6	5.1	975	2	US-09-270-767-42040	Sequence 42040, A	1201	5	4.2	5	2	US-09-688-678-8	Sequence 8, Appli
1129	6	5.1	986	1	US-08-548-159-3	Sequence 3, Appli	1202	5	4.2	5	2	US-09-790-497A-437	Sequence 437, App
1130	6	5.1	998	2	US-09-252-991A-28424	Sequence 28424, A	1203	5	4.2	5	2	US-09-790-497A-532	Sequence 532, App
1131	6	5.1	1012	2	US-08-811-481-16	Sequence 16, Appl	1204	5	4.2	5	2	US-09-576-824A-437	Sequence 437, App
1132	6	5.1	1012	2	US-09-876-527-16	Sequence 16, Appl	1205	5	4.2	5	2	US-08-925-002-25	Sequence 25, Appl
1133	6	5.1	1015	2	US-09-344-510B-5	Sequence 5, Appli	1206	5	4.2	5	2	US-09-910-552-25	Sequence 25, Appl
1134	6	5.1	1015	2	US-09-949-016-6276	Sequence 6276, Ap	1207	5	4.2	5	2	5194596-23	Patent No. 5194596
1135	6	5.1	1017	2	US-09-762-724-12	Sequence 12, Appl	1208	5	4.2	5	2	5219739-28	Patent No. 5219739
1136	6	5.1	1023	2	US-09-762-724-14	Sequence 14, Appl	1209	5	4.2	5	2	US-08-102-738-15	Sequence 15, Appl
1137	6	5.1	1027	2	US-09-252-991A-26216	Sequence 26216, A	1210	5	4.2	5	2	US-08-102-738-16	Sequence 16, Appl
1138	6	5.1	1037	2	US-09-428-711A-21	Sequence 21, Appl	1211	5	4.2	5	2	US-08-615-279-14	Sequence 14, Appl
1139	6	5.1	1043	2	US-09-252-991A-32210	Sequence 32210, A	1212	5	4.2	5	2	US-08-615-279-36	Sequence 36, Appl
1140	6	5.1	1044	2	US-09-252-991A-24495	Sequence 24495, A	1213	5	4.2	5	2	US-08-146-028-262	Sequence 262, App
1141	6	5.1	1047	2	US-09-543-681A-6186	Sequence 6186, Ap	1214	5	4.2	5	2	US-08-146-028-263	Sequence 263, App
1142	6	5.1	1150	2	US-09-902-540-13969	Sequence 13969, A	1215	5	4.2	5	2	US-08-146-028-264	Sequence 264, App
1143	6	5.1	1156	2	US-09-198-452A-171	Sequence 171, App	1216	5	4.2	5	2	US-08-146-028-282	Sequence 282, App
1144	6	5.1	1224	2	US-09-252-991A-32880	Sequence 32880, A	1217	5	4.2	5	2	US-08-146-028-288	Sequence 288, App
1145	6	5.1	1252	2	US-09-902-540-13967	Sequence 13967, A	1218	5	4.2	5	2	US-08-146-028-285	Sequence 285, App
1146	6	5.1	1261	1	US-08-764-100-26	Sequence 26, Appl	1219	5	4.2	5	2	US-08-146-028-301	Sequence 301, App
1147	6	5.1	1274	2	US-09-602-777A-360	Sequence 360, App	1220	5	4.2	5	2	US-08-146-028-304	Sequence 304, App
1148	6	5.1	1284	2	US-09-343-494-9	Sequence 9, Appli	1221	5	4.2	5	2	US-08-146-028-305	Sequence 305, App
1149	6	5.1	1284	2	US-09-358-383C-11	Sequence 11, Appl	1222	5	4.2	5	2	US-08-146-028-306	Sequence 306, App
1150	6	5.1	1284	2	US-10-160-224-9	Sequence 9, Appli	1223	5	4.2	5	2	US-08-146-028-307	Sequence 307, App
1151	6	5.1	1318	2	US-10-237-551-197	Sequence 197, App	1224	5	4.2	5	2	US-08-146-028-308	Sequence 308, App
1152	6	5.1	1351	2	US-09-489-039A-11032	Sequence 11032, A	1225	5	4.2	5	2	US-08-146-028-310	Sequence 310, App
1153	6	5.1	1372	2	US-09-902-540-14099	Sequence 14099, A	1226	5	4.2	5	2	US-08-146-028-311	Sequence 311, App
1154	6	5.1	1380	2	US-09-252-991A-25722	Sequence 25722, A	1227	5	4.2	5	2	US-08-146-028-312	Sequence 312, App
1155	6	5.1	1399	2	US-08-462-467B-14	Sequence 14, Appl	1228	5	4.2	5	2	US-08-146-028-313	Sequence 313, App
1156	6	5.1	1409	2	US-09-489-039A-10338	Sequence 10338, A	1229	5	4.2	5	2	US-08-146-028-314	Sequence 314, App
1157	6	5.1	1411	2	US-09-252-991A-23628	Sequence 23628, A	1230	5	4.2	5	2	US-08-146-028-333	Sequence 333, App
1158	6	5.1	1444	2	US-09-949-016-9652	Sequence 9652, Ap	1231	5	4.2	5	2	US-08-146-028-339	Sequence 339, App
1159	6	5.1	1461	1	US-08-993-228-10	Sequence 10, Appl	1232	5	4.2	5	2	US-08-146-028-340	Sequence 340, App
1160	6	5.1	1495	2	US-08-462-467B-12	Sequence 12, Appl	1233	5	4.2	5	2	US-08-146-028-341	Sequence 341, App
1161	6	5.1	1562	2	US-09-438-185A-152	Sequence 152, App	1234	5	4.2	5	2	US-08-146-028-342	Sequence 342, App
1162	6	5.1	1618	2	US-08-462-467B-4	Sequence 4, Appli	1235	5	4.2	5	2	US-08-146-028-343	Sequence 343, App
1163	6	5.1	1627	2	US-09-252-991A-28863	Sequence 28863, A	1236	5	4.2	5	2	US-08-146-028-346	Sequence 346, App
1164	6	5.1	1728	2	US-10-037-417-72	Sequence 72, Appl	1237	5	4.2	5	2	US-08-146-028-347	Sequence 347, App
1165	6	5.1	1765	2	US-09-354-147C-2	Sequence 2, Appli	1238	5	4.2	5	2	US-08-146-028-348	Sequence 348, App
1166	6	5.1	1765	2	US-09-354-147C-3	Sequence 3, Appli	1239	5	4.2	5	2	US-08-146-028-349	Sequence 349, App
1167	6	5.1	1765	2	US-09-354-147C-5	Sequence 5, Appli	1240	5	4.2	5	2	US-08-845-926-37	Sequence 37, Appl
1168	6	5.1	1791	2	US-09-354-147C-42	Sequence 42, Appl	1241	5	4.2	5	2	US-08-845-926-38	Sequence 38, Appl
1169	6	5.1	1891	1	US-08-804-227C-12	Sequence 12, Appl	1242	5	4.2	5	2	US-08-845-926-41	Sequence 41, Appl
1170	6	5.1	1891	1	US-08-804-198-6	Sequence 6, Appli	1243	5	4.2	5	2	US-08-803-981-213	Sequence 213, App
1171	6	5.1	2004	2	US-09-538-092-1371	Sequence 1371, Ap	1244	5	4.2	5	2	US-08-723-425A-262	Sequence 262, App
1172	6	5.1	2004	2	US-09-949-016-6756	Sequence 6756, Ap	1245	5	4.2	5	2	US-08-723-425A-263	Sequence 263, App
1173	6	5.1	2472	2	US-09-252-991A-17052	Sequence 17052, A	1246	5	4.2	5	2	US-08-723-425A-264	Sequence 264, App
1174	6	5.1	2516	2	US-08-374-077C-2	Sequence 2, Appli	1247	5	4.2	5	2	US-08-723-425A-282	Sequence 282, App
1175	6	5.1	2516	2	US-08-895-590-2	Sequence 2, Appli	1248	5	4.2	5	2	US-08-723-425A-288	Sequence 288, App
1176	6	5.1	2516	2	US-09-539-879A-2	Sequence 2, Appli	1249	5	4.2	5	2	US-08-723-425A-295	Sequence 295, App
1177	6	5.1	2887	2	US-08-462-467B-8	Sequence 8, Appli	1250	5	4.2	5	2	US-08-723-425A-301	Sequence 301, App
1178	6	5.1	2887	2	US-08-462-467B-8	Sequence 8, Appli	1251	5	4.2	5	2	US-08-723-425A-304	Sequence 304, App
1179	6	5.1	2987	1	US-08-970-269A-29	Sequence 29, Appl	1252	5	4.2	5	2	US-08-723-425A-305	Sequence 305, App
1180	6	5.1	2987	1	US-09-407-562-29	Sequence 29, Appl	1253	5	4.2	5	2	US-08-723-425A-306	Sequence 306, App
1181	6	5.1	3033	1	US-07-925-695-8	Sequence 8, Appli	1254	5	4.2	5	2	US-08-723-425A-307	Sequence 307, App
1182	6	5.1	3033	1	US-07-925-695-9	Sequence 9, Appli	1255	5	4.2	5	2	US-08-723-425A-308	Sequence 308, App
1183	6	5.1	3168	2	US-09-489-039A-14067	Sequence 14067, A	1256	5	4.2	5	2	US-08-723-425A-310	Sequence 310, App
1184	6	5.1	3170	2	US-09-036-987A-4	Sequence 4, Appli	1257	5	4.2	5	2	US-08-723-425A-311	Sequence 311, App
1185	6	5.1	3170	2	US-09-370-700-4	Sequence 4, Appli	1258	5	4.2	5	2	US-08-723-425A-312	Sequence 312, App
1186	6	5.1	3170	2	US-06-603-207-4	Sequence 4, Appli	1259	5	4.2	5	2	US-08-723-425A-313	Sequence 313, App
1187	6	5.1	3724	1	US-08-804-227C-10	Sequence 10, Appl	1260	5	4.2	5	2	US-08-723-425A-314	Sequence 314, App
1188	6	5.1	3724	1	US-08-804-198-4	Sequence 4, Appli	1261	5	4.2	5	2	US-08-723-425A-333	Sequence 333, App
1189	6	5.1	3959	1	US-08-970-269A-30	Sequence 30, Appl	1262	5	4.2	5	2	US-08-723-425A-339	Sequence 339, App
1190	6	5.1	3959	1	US-09-407-562-30	Sequence 30, Appl	1263	5	4.2	5	2	US-08-723-425A-340	Sequence 340, App
1191	6	5.1	4096	2	US-09-296-662-34	Sequence 34, Appl	1264	5	4.2	5	2	US-08-723-425A-341	Sequence 341, App
1192	6	5.1	4928	2	US-09-036-987A-5	Sequence 5, Appli	1265	5	4.2	5	2	US-08-723-425A-342	Sequence 342, App
1193	6	5.1	4928	2	US-09-370-700-5	Sequence 5, Appli	1266	5	4.2	5	2	US-08-723-425A-343	Sequence 343, App
1194	6	5.1	4928	2	US-09-603-207-5	Sequence 5, Appli	1267	5	4.2	5	2	US-08-723-425A-346	Sequence 346, App
1195	5	4.2	5	1	US-08-350-260A-411	Sequence 411, App	1268	5	4.2	5	2		

1269	5	4.2	9	2	US-08-723-425A-347	Sequence 347, App	1342	5	4.2	9	2	US-09-576-824A-283	Sequence 283, App
1270	5	4.2	9	2	US-08-723-425A-348	Sequence 348, App	1343	5	4.2	9	2	US-09-576-824A-284	Sequence 284, App
1271	5	4.2	9	2	US-08-723-425A-349	Sequence 349, App	1344	5	4.2	9	2	US-09-576-824A-285	Sequence 285, App
1272	5	4.2	9	2	US-08-604-365-7	Sequence 7, Appli	1345	5	4.2	9	2	US-09-576-824A-286	Sequence 286, App
1273	5	4.2	9	2	US-08-604-365-9	Sequence 9, Appli	1346	5	4.2	9	2	US-09-576-824A-299	Sequence 299, App
1274	5	4.2	9	2	US-09-112-206-262	Sequence 262, App	1347	5	4.2	9	2	US-09-576-824A-300	Sequence 300, App
1275	5	4.2	9	2	US-09-112-206-263	Sequence 263, App	1348	5	4.2	9	2	US-09-576-824A-301	Sequence 301, App
1276	5	4.2	9	2	US-09-112-206-264	Sequence 264, App	1349	5	4.2	9	2	US-09-576-824A-302	Sequence 302, App
1277	5	4.2	9	2	US-09-112-206-282	Sequence 282, App	1350	5	4.2	9	2	US-09-576-824A-303	Sequence 303, App
1278	5	4.2	9	2	US-09-112-206-288	Sequence 288, App	1351	5	4.2	9	2	US-09-576-824A-592	Sequence 592, App
1279	5	4.2	9	2	US-09-112-206-295	Sequence 295, App	1352	5	4.2	9	2	US-09-680-497-262	Sequence 262, App
1280	5	4.2	9	2	US-09-112-206-301	Sequence 301, App	1353	5	4.2	9	2	US-09-680-497-263	Sequence 263, App
1281	5	4.2	9	2	US-09-112-206-304	Sequence 304, App	1354	5	4.2	9	2	US-09-680-497-264	Sequence 264, App
1282	5	4.2	9	2	US-09-112-206-305	Sequence 305, App	1355	5	4.2	9	2	US-09-680-497-282	Sequence 282, App
1283	5	4.2	9	2	US-09-112-206-306	Sequence 306, App	1356	5	4.2	9	2	US-09-680-497-288	Sequence 288, App
1284	5	4.2	9	2	US-09-112-206-307	Sequence 307, App	1357	5	4.2	9	2	US-09-680-497-295	Sequence 295, App
1285	5	4.2	9	2	US-09-112-206-308	Sequence 308, App	1358	5	4.2	9	2	US-09-680-497-301	Sequence 301, App
1286	5	4.2	9	2	US-09-112-206-310	Sequence 310, App	1359	5	4.2	9	2	US-09-680-497-304	Sequence 304, App
1287	5	4.2	9	2	US-09-112-206-311	Sequence 311, App	1360	5	4.2	9	2	US-09-680-497-305	Sequence 305, App
1288	5	4.2	9	2	US-09-112-206-312	Sequence 312, App	1361	5	4.2	9	2	US-09-680-497-306	Sequence 306, App
1289	5	4.2	9	2	US-09-112-206-313	Sequence 313, App	1362	5	4.2	9	2	US-09-680-497-307	Sequence 307, App
1290	5	4.2	9	2	US-09-112-206-314	Sequence 314, App	1363	5	4.2	9	2	US-09-680-497-308	Sequence 308, App
1291	5	4.2	9	2	US-09-112-206-333	Sequence 333, App	1364	5	4.2	9	2	US-09-680-497-310	Sequence 310, App
1292	5	4.2	9	2	US-09-112-206-339	Sequence 339, App	1365	5	4.2	9	2	US-09-680-497-311	Sequence 311, App
1293	5	4.2	9	2	US-09-112-206-340	Sequence 340, App	1366	5	4.2	9	2	US-09-680-497-312	Sequence 312, App
1294	5	4.2	9	2	US-09-112-206-341	Sequence 341, App	1367	5	4.2	9	2	US-09-680-497-313	Sequence 313, App
1295	5	4.2	9	2	US-09-112-206-342	Sequence 342, App	1368	5	4.2	9	2	US-09-680-497-314	Sequence 314, App
1296	5	4.2	9	2	US-09-112-206-343	Sequence 343, App	1369	5	4.2	9	2	US-09-680-497-333	Sequence 333, App
1297	5	4.2	9	2	US-09-112-206-346	Sequence 346, App	1370	5	4.2	9	2	US-09-680-497-339	Sequence 339, App
1298	5	4.2	9	2	US-09-112-206-347	Sequence 347, App	1371	5	4.2	9	2	US-09-680-497-340	Sequence 340, App
1299	5	4.2	9	2	US-09-112-206-348	Sequence 348, App	1372	5	4.2	9	2	US-09-680-497-341	Sequence 341, App
1300	5	4.2	9	2	US-09-112-206-349	Sequence 349, App	1373	5	4.2	9	2	US-09-680-497-342	Sequence 342, App
1301	5	4.2	9	2	US-09-351-296-37	Sequence 37, Appl	1374	5	4.2	9	2	US-09-680-497-343	Sequence 343, App
1302	5	4.2	9	2	US-09-351-296-38	Sequence 38, Appl	1375	5	4.2	9	2	US-09-680-497-346	Sequence 346, App
1303	5	4.2	9	2	US-09-351-296-41	Sequence 41, Appl	1376	5	4.2	9	2	US-09-680-497-347	Sequence 347, App
1304	5	4.2	9	2	US-08-776-189-14	Sequence 14, Appl	1377	5	4.2	9	2	US-09-680-497-348	Sequence 348, App
1305	5	4.2	9	2	US-08-776-189-29	Sequence 29, Appl	1378	5	4.2	9	2	US-09-680-497-349	Sequence 349, App
1306	5	4.2	9	2	US-09-689-678-7	Sequence 7, Appli	1379	5	4.2	9	2	US-09-453-174-14	Sequence 14, Appl
1307	5	4.2	9	2	US-09-689-678-9	Sequence 9, Appli	1380	5	4.2	9	2	US-09-453-174-29	Sequence 29, Appl
1308	5	4.2	9	2	US-08-776-188C-14	Sequence 14, Appl	1381	5	4.2	9	6	5310729-41	Patent No. 5310729
1309	5	4.2	9	2	US-09-790-497A-258	Sequence 258, App	1382	5	4.2	10	1	US-08-214-650-32	Sequence 32, Appl
1310	5	4.2	9	2	US-09-790-497A-259	Sequence 259, App	1383	5	4.2	10	1	US-08-617-929-15	Sequence 15, Appl
1311	5	4.2	9	2	US-09-790-497A-260	Sequence 260, App	1384	5	4.2	10	1	US-08-617-929-19	Sequence 19, Appl
1312	5	4.2	9	2	US-09-790-497A-272	Sequence 272, App	1385	5	4.2	10	1	US-08-617-929-34	Sequence 34, Appl
1313	5	4.2	9	2	US-09-790-497A-279	Sequence 279, App	1386	5	4.2	10	2	US-08-822-586-1	Sequence 1, Appli
1314	5	4.2	9	2	US-09-790-497A-282	Sequence 282, App	1387	5	4.2	10	2	US-08-604-365-10	Sequence 10, Appl
1315	5	4.2	9	2	US-09-790-497A-283	Sequence 283, App	1388	5	4.2	10	2	US-09-011-961-18	Sequence 18, Appl
1316	5	4.2	9	2	US-09-790-497A-284	Sequence 284, App	1389	5	4.2	10	2	US-09-011-961-20	Sequence 20, Appl
1317	5	4.2	9	2	US-09-790-497A-285	Sequence 285, App	1390	5	4.2	10	2	US-09-011-961-23	Sequence 23, Appl
1318	5	4.2	9	2	US-09-790-497A-286	Sequence 286, App	1391	5	4.2	10	2	US-09-011-961-29	Sequence 29, Appl
1319	5	4.2	9	2	US-09-790-497A-299	Sequence 299, App	1392	5	4.2	10	2	US-09-011-961-30	Sequence 30, Appl
1320	5	4.2	9	2	US-09-790-497A-300	Sequence 300, App	1393	5	4.2	10	2	US-09-011-961-31	Sequence 31, Appl
1321	5	4.2	9	2	US-09-790-497A-301	Sequence 301, App	1394	5	4.2	10	2	US-09-011-961-32	Sequence 32, Appl
1322	5	4.2	9	2	US-09-790-497A-302	Sequence 302, App	1395	5	4.2	10	2	US-09-011-961-33	Sequence 33, Appl
1323	5	4.2	9	2	US-09-790-497A-303	Sequence 303, App	1396	5	4.2	10	2	US-09-011-961-34	Sequence 34, Appl
1324	5	4.2	9	2	US-09-790-497A-515	Sequence 515, App	1397	5	4.2	10	2	US-09-011-961-36	Sequence 36, Appl
1325	5	4.2	9	2	US-09-790-497A-516	Sequence 516, App	1398	5	4.2	10	2	US-09-689-678-10	Sequence 10, Appl
1326	5	4.2	9	2	US-09-790-497A-524	Sequence 524, App	1399	5	4.2	10	2	US-09-755-630B-231	Sequence 231, App
1327	5	4.2	9	2	US-09-790-497A-525	Sequence 525, App	1400	5	4.2	10	2	US-09-755-630B-224	Sequence 224, App
1328	5	4.2	9	2	US-09-790-497A-526	Sequence 526, App	1401	5	4.2	10	2	US-10-658-180-221	Sequence 221, App
1329	5	4.2	9	2	US-09-790-497A-527	Sequence 527, App	1402	5	4.2	10	2	US-10-658-180-234	Sequence 234, App
1330	5	4.2	9	2	US-09-790-497A-528	Sequence 528, App	1403	5	4.2	11	1	US-07-666-719-20	Sequence 20, Appl
1331	5	4.2	9	2	US-09-790-497A-538	Sequence 538, App	1404	5	4.2	11	1	US-07-699-468-1	Sequence 1, Appli
1332	5	4.2	9	2	US-09-790-497A-539	Sequence 539, App	1405	5	4.2	11	1	US-07-699-468-2	Sequence 2, Appli
1333	5	4.2	9	2	US-09-790-497A-540	Sequence 540, App	1406	5	4.2	11	1	US-07-699-468-3	Sequence 3, Appli
1334	5	4.2	9	2	US-09-790-497A-541	Sequence 541, App	1407	5	4.2	11	1	US-07-699-468-4	Sequence 4, Appli
1335	5	4.2	9	2	US-09-790-497A-542	Sequence 542, App	1408	5	4.2	11	1	US-07-699-468-5	Sequence 5, Appli
1336	5	4.2	9	2	US-09-576-824A-258	Sequence 258, App	1409	5	4.2	11	1	US-07-723-002C-14	Sequence 14, Appl
1337	5	4.2	9	2	US-09-576-824A-259	Sequence 259, App	1410	5	4.2	11	2	US-09-478-479B-26	Sequence 26, Appl
1338	5	4.2	9	2	US-09-576-824A-260	Sequence 260, App	1411	5	4.2	11	2	US-09-576-824A-515	Sequence 515, App
1339	5	4.2	9	2	US-09-576-824A-272	Sequence 272, App	1412	5	4.2	11	2	US-09-576-824A-516	Sequence 516, App
1340	5	4.2	9	2	US-09-576-824A-279	Sequence 279, App	1413	5	4.2	11	2	US-09-576-824A-524	Sequence 524, App
1341	5	4.2	9	2	US-09-576-824A-282	Sequence 282, App	1414	5	4.2	11	2	US-09-576-824A-525	Sequence 525, App

1415	5	4.2	11	2	US-09-576-824A-526	Sequence 526, App	1488	14	1	US-08-232-453A-29	Sequence 29, Appl
1416	5	4.2	11	2	US-09-576-824A-527	Sequence 527, App	1489	14	1	US-08-232-453A-30	Sequence 30, Appl
1417	5	4.2	11	2	US-09-576-824A-528	Sequence 528, App	1490	14	1	US-08-232-453A-31	Sequence 31, Appl
1418	5	4.2	11	2	US-09-576-824A-529	Sequence 529, App	1491	14	1	US-08-232-453A-32	Sequence 32, Appl
1419	5	4.2	11	2	US-09-576-824A-530	Sequence 530, App	1492	14	1	US-08-232-453A-33	Sequence 33, Appl
1420	5	4.2	11	2	US-09-576-824A-531	Sequence 531, App	1493	14	1	US-08-232-453A-34	Sequence 34, Appl
1421	5	4.2	11	2	US-09-576-824A-541	Sequence 541, App	1494	14	1	US-08-232-453A-37	Sequence 37, Appl
1422	5	4.2	11	2	US-09-576-824A-542	Sequence 542, App	1495	14	1	US-08-232-453A-38	Sequence 38, Appl
1423	5	4.2	11	2	US-09-390-280-1	Sequence 1, Appl	1496	14	1	US-08-232-453A-39	Sequence 39, Appl
1424	5	4.2	11	2	US-09-390-280-2	Sequence 2, Appl	1497	14	1	US-08-232-453A-40	Sequence 40, Appl
1425	5	4.2	11	2	US-09-390-280-3	Sequence 3, Appl	1498	14	1	US-08-232-453A-46	Sequence 46, Appl
1426	5	4.2	11	2	US-09-390-280-4	Sequence 4, Appl	1499	14	1	US-08-232-453A-66	Sequence 66, Appl
1427	5	4.2	11	2	US-09-390-280-5	Sequence 5, Appl	1500	14	1	US-08-156-552A-1	Sequence 1, Appl
1428	5	4.2	11	2	US-09-390-280-6	Sequence 6, Appl	1501	14	1	US-08-156-552A-2	Sequence 2, Appl
1429	5	4.2	11	2	US-09-650-684-1	Sequence 1, Appl	1502	14	1	US-08-818-552-33	Sequence 33, Appl
1430	5	4.2	11	2	US-09-650-684-2	Sequence 2, Appl	1503	14	2	US-08-116-294-1	Sequence 1, Appl
1431	5	4.2	11	2	US-09-650-684-3	Sequence 3, Appl	1504	14	2	US-08-782-997A-6	Sequence 6, Appl
1432	5	4.2	11	2	US-09-650-684-4	Sequence 4, Appl	1505	14	2	US-08-641-873-6	Sequence 6, Appl
1433	5	4.2	11	2	US-09-650-684-5	Sequence 5, Appl	1506	14	2	US-09-035-249A-1	Sequence 1, Appl
1434	5	4.2	11	2	US-09-650-684-6	Sequence 6, Appl	1507	14	2	US-09-035-249A-2	Sequence 2, Appl
1435	5	4.2	11	2	US-09-650-684-7	Sequence 7, Appl	1508	14	2	US-09-058-562-28	Sequence 28, Appl
1436	5	4.2	11	2	US-09-650-684-8	Sequence 8, Appl	1509	14	2	US-09-011-961-6	Sequence 6, Appl
1437	5	4.2	11	6	US-09-965-594A-26	Sequence 26, Appl	1510	14	2	US-08-818-252-33	Sequence 33, Appl
1438	5	4.2	11	6	5169773-1	Patent No. 5169773	1511	14	2	US-09-041-886-51	Sequence 51, Appl
1439	5	4.2	12	1	5202310-1	Patent No. 5202310	1512	14	2	US-08-842-322-27	Sequence 27, Appl
1440	5	4.2	12	1	US-07-789-179-3	Sequence 3, Appl	1513	14	2	US-08-839-131A-5	Sequence 5, Appl
1441	5	4.2	12	1	US-08-240-712-30	Sequence 30, Appl	1514	14	2	US-09-316-919-49	Sequence 49, Appl
1442	5	4.2	12	1	US-08-250-789A-30	Sequence 30, Appl	1515	14	2	US-09-428-082B-176	Sequence 176, App
1443	5	4.2	12	1	US-08-250-789A-47	Sequence 47, Appl	1516	14	2	US-09-430-619-215	Sequence 215, App
1444	5	4.2	12	1	US-08-156-552A-3	Sequence 3, Appl	1517	14	2	US-09-428-082B-1107	Sequence 1107, App
1445	5	4.2	12	1	US-08-443-890-30	Sequence 30, Appl	1518	14	2	US-09-316-920A-49	Sequence 49, Appl
1446	5	4.2	12	1	US-08-444-942-3	Sequence 3, Appl	1519	14	2	US-08-453-485B-3	Sequence 3, Appl
1447	5	4.2	12	1	US-08-446-105-3	Sequence 3, Appl	1520	14	2	US-09-623-548A-1411	Sequence 1411, Ap
1448	5	4.2	12	1	US-08-444-939-3	Sequence 3, Appl	1521	14	2	US-09-623-548A-1412	Sequence 1412, Ap
1449	5	4.2	12	1	US-08-615-279-13	Sequence 13, Appl	1522	14	2	US-09-623-548A-1413	Sequence 1413, Ap
1450	5	4.2	12	1	US-08-615-279-35	Sequence 35, Appl	1523	14	2	US-09-623-548A-1513	Sequence 1513, Ap
1451	5	4.2	12	1	US-08-444-991-3	Sequence 3, Appl	1524	14	2	US-09-841-091B-25	Sequence 25, Appl
1452	5	4.2	12	1	US-08-450-733-3	Sequence 3, Appl	1525	14	2	US-09-657-276-1411	Sequence 1411, Ap
1453	5	4.2	12	1	US-08-617-929-37	Sequence 37, Appl	1526	14	2	US-09-657-276-1412	Sequence 1412, Ap
1454	5	4.2	12	1	US-08-617-929-42	Sequence 42, Appl	1527	14	2	US-09-657-276-1413	Sequence 1413, Ap
1455	5	4.2	12	2	US-09-035-249A-3	Sequence 3, Appl	1528	14	2	US-09-657-276-1513	Sequence 1513, Ap
1456	5	4.2	12	2	US-09-011-961-14	Sequence 14, Appl	1529	15	2	US-09-444-281-96	Sequence 96, Appl
1457	5	4.2	12	2	US-09-011-961-17	Sequence 17, Appl	1530	15	2	US-08-604-365-3	Sequence 3, Appl
1458	5	4.2	12	2	US-08-344-915A-3	Sequence 3, Appl	1531	15	2	US-09-009-953-28	Sequence 28, Appl
1459	5	4.2	12	2	US-09-344-456-2	Sequence 2, Appl	1532	15	2	US-09-009-953-69	Sequence 69, Appl
1460	5	4.2	12	2	US-08-776-188C-13	Sequence 13, Appl	1533	15	2	US-08-776-189-25	Sequence 25, Appl
1461	5	4.2	12	2	US-08-776-188C-59	Sequence 59, Appl	1534	15	2	US-09-689-678-3	Sequence 3, Appl
1462	5	4.2	12	2	US-09-453-174-13	Sequence 13, Appl	1535	15	6	US-09-453-174-25	Sequence 25, Appl
1463	5	4.2	12	2	US-09-551-151A-75	Sequence 75, Appl	1536	16	1	5223254-3	Patent No. 5223254
1464	5	4.2	12	2	US-09-559-021-26	Sequence 26, Appl	1537	16	1	US-08-096-946-3	Sequence 3, Appl
1465	5	4.2	12	2	US-08-632-514C-25	Sequence 25, Appl	1538	16	2	US-08-318-200-11	Sequence 11, Appl
1466	5	4.2	13	1	US-09-188-177-25	Sequence 25, Appl	1539	16	2	US-09-889-136-9	Sequence 9, Appl
1467	5	4.2	13	2	US-09-011-961-47	Sequence 47, Appl	1540	16	4	US-08-776-188C-58	Sequence 58, Appl
1468	5	4.2	13	2	US-08-776-188C-74	Sequence 74, Appl	1541	16	4	PCT-US91-02942-14	Sequence 14, Appl
1469	5	4.2	13	2	US-09-890-065-4	Sequence 4, Appl	1542	17	1	US-08-441-871-116	Sequence 3, Appl
1470	5	4.2	14	1	US-07-961-837-6	Sequence 6, Appl	1543	17	2	US-08-641-873-14	Sequence 14, Appl
1471	5	4.2	14	1	US-08-232-453A-1	Sequence 1, Appl	1544	17	2	US-08-983-075D-14	Sequence 14, Appl
1472	5	4.2	14	1	US-08-232-453A-2	Sequence 2, Appl	1545	17	2	US-09-244-583-10	Sequence 10, Appl
1473	5	4.2	14	1	US-08-232-453A-3	Sequence 3, Appl	1546	17	2	US-09-368-670C-5	Sequence 5, Appl
1474	5	4.2	14	1	US-08-232-453A-4	Sequence 4, Appl	1547	18	1	US-09-719-261B-8	Sequence 8, Appl
1475	5	4.2	14	1	US-08-232-453A-5	Sequence 5, Appl	1548	18	1	US-07-666-719-12	Sequence 12, Appl
1476	5	4.2	14	1	US-08-232-453A-6	Sequence 6, Appl	1549	18	1	US-08-244-116B-9	Sequence 9, Appl
1477	5	4.2	14	1	US-08-232-453A-8	Sequence 8, Appl	1550	18	1	US-08-283-917-14	Sequence 14, Appl
1478	5	4.2	14	1	US-08-232-453A-16	Sequence 16, Appl	1551	18	1	US-08-686-594-11	Sequence 11, Appl
1479	5	4.2	14	1	US-08-232-453A-18	Sequence 18, Appl	1552	18	1	US-08-961-716-14	Sequence 14, Appl
1480	5	4.2	14	1	US-08-232-453A-19	Sequence 19, Appl	1553	18	1	US-09-017-205-16	Sequence 16, Appl
1481	5	4.2	14	1	US-08-232-453A-20	Sequence 20, Appl	1554	18	1	US-09-017-205-17	Sequence 17, Appl
1482	5	4.2	14	1	US-08-232-453A-22	Sequence 22, Appl	1555	18	2	US-08-802-981-97	Sequence 97, Appl
1483	5	4.2	14	1	US-08-232-453A-23	Sequence 23, Appl	1556	18	2	US-08-604-365-1	Sequence 1, Appl
1484	5	4.2	14	1	US-08-232-453A-24	Sequence 24, Appl	1557	18	2	US-08-604-365-4	Sequence 4, Appl
1485	5	4.2	14	1	US-08-232-453A-27	Sequence 27, Appl	1558	18	2	US-09-268-480-22	Sequence 22, Appl
1486	5	4.2	14	1	US-08-232-453A-28	Sequence 28, Appl	1559	18	2	US-09-689-678-1	Sequence 1, Appl
1487	5	4.2	14	1	US-08-232-453A-29	Sequence 29, Appl	1560	18	2	US-09-755-100A-15	Sequence 15, Appl

1561	5	4.2	18	2	US-09-471-276-1120	Sequence 1120, Ap	1634	5	4.2	20	2	US-09-878-281A-274	Sequence 274, App
1562	5	4.2	18	2	US-08-537-802-24	Sequence 24, Appl	1635	5	4.2	20	2	US-09-878-281A-275	Sequence 275, App
1563	5	4.2	18	2	US-08-537-802-26	Sequence 26, Appl	1636	5	4.2	20	2	US-09-878-281A-276	Sequence 276, App
1564	5	4.2	18	2	US-09-747-287A-140	Sequence 140, App	1637	5	4.2	20	2	US-09-878-281A-277	Sequence 277, App
1565	5	4.2	18	2	US-09-394-019C-93	Sequence 93, Appl	1638	5	4.2	20	2	US-10-044-995-9	Sequence 9, Appl
1566	5	4.2	18	2	US-09-394-019C-338	Sequence 338, App	1639	5	4.2	20	2	US-10-044-995-12	Sequence 12, Appl
1567	5	4.2	19	1	US-08-686-594-12	Sequence 12, Appl	1640	5	4.2	20	2	US-10-044-995-13	Sequence 13, Appl
1568	5	4.2	19	2	US-08-802-981-95	Sequence 95, Appl	1641	5	4.2	20	2	US-10-044-995-14	Sequence 14, Appl
1569	5	4.2	19	2	US-08-802-981-100	Sequence 100, App	1642	5	4.2	20	2	US-10-044-995-15	Sequence 15, Appl
1570	5	4.2	19	2	US-08-537-802-38	Sequence 38, Appl	1643	5	4.2	20	2	US-09-747-287A-146	Sequence 146, App
1571	5	4.2	19	2	US-08-537-802-40	Sequence 40, App	1644	5	4.2	20	2	US-09-394-019C-99	Sequence 99, Appl
1572	5	4.2	19	2	US-09-747-287A-138	Sequence 138, App	1645	5	4.2	20	2	US-09-394-019C-344	Sequence 344, App
1573	5	4.2	19	2	US-09-747-287A-143	Sequence 143, App	1646	5	4.2	21	1	US-08-127-499A-33	Sequence 33, Appl
1574	5	4.2	19	2	US-09-394-019C-91	Sequence 91, Appl	1647	5	4.2	21	1	US-08-482-847-33	Sequence 33, Appl
1575	5	4.2	19	2	US-09-394-019C-96	Sequence 96, Appl	1648	5	4.2	21	1	US-08-441-871-83	Sequence 83, Appl
1576	5	4.2	19	2	US-09-394-019C-336	Sequence 336, App	1649	5	4.2	21	2	US-08-604-365-5	Sequence 5, Appl
1577	5	4.2	19	2	US-09-394-019C-341	Sequence 341, App	1650	5	4.2	21	2	US-09-689-678-5	Sequence 5, Appl
1578	5	4.2	20	1	US-08-787-547-5	Sequence 5, Appl	1651	5	4.2	21	2	US-09-409-604-17	Sequence 17, Appl
1579	5	4.2	20	1	US-08-686-594-13	Sequence 13, Appl	1652	5	4.2	21	2	US-09-962-756-701	Sequence 701, App
1580	5	4.2	20	1	US-08-466-975A-9	Sequence 9, Appl	1653	5	4.2	21	6	5168050-1	Patent No. 5168050
1581	5	4.2	20	1	US-08-466-975A-12	Sequence 12, Appl	1654	5	4.2	22	1	US-08-146-028-49	Sequence 49, Appl
1582	5	4.2	20	1	US-08-466-975A-13	Sequence 13, Appl	1655	5	4.2	22	1	US-08-146-028-53	Sequence 53, Appl
1583	5	4.2	20	1	US-08-466-975A-14	Sequence 14, Appl	1656	5	4.2	22	1	US-08-146-028-54	Sequence 54, Appl
1584	5	4.2	20	1	US-08-466-975A-15	Sequence 15, Appl	1657	5	4.2	22	1	US-08-146-028-55	Sequence 55, Appl
1585	5	4.2	20	1	US-08-391-671A-9	Sequence 9, Appl	1658	5	4.2	22	1	US-08-146-028-56	Sequence 56, Appl
1586	5	4.2	20	1	US-08-391-671A-12	Sequence 12, Appl	1659	5	4.2	22	1	US-08-146-028-116	Sequence 116, App
1587	5	4.2	20	1	US-08-391-671A-13	Sequence 13, Appl	1660	5	4.2	22	1	US-08-146-028-142	Sequence 142, App
1588	5	4.2	20	1	US-08-391-671A-14	Sequence 14, Appl	1661	5	4.2	22	1	US-08-146-028-143	Sequence 143, App
1589	5	4.2	20	1	US-08-391-671A-15	Sequence 15, Appl	1662	5	4.2	22	1	US-08-146-028-148	Sequence 148, App
1590	5	4.2	20	1	US-08-078-311-6	Sequence 6, Appl	1663	5	4.2	22	1	US-08-146-028-150	Sequence 150, App
1591	5	4.2	20	1	US-08-460-402-6	Sequence 6, Appl	1664	5	4.2	22	1	US-08-146-028-152	Sequence 152, App
1592	5	4.2	20	2	US-08-467-902A-9	Sequence 9, Appl	1665	5	4.2	22	2	US-08-641-873-9	Sequence 9, Appl
1593	5	4.2	20	2	US-08-467-902A-12	Sequence 12, Appl	1666	5	4.2	22	2	US-08-723-425A-49	Sequence 49, Appl
1594	5	4.2	20	2	US-08-467-902A-13	Sequence 13, Appl	1667	5	4.2	22	2	US-08-723-425A-53	Sequence 53, Appl
1595	5	4.2	20	2	US-08-467-902A-14	Sequence 14, Appl	1668	5	4.2	22	2	US-08-723-425A-54	Sequence 54, Appl
1596	5	4.2	20	2	US-08-467-902A-15	Sequence 15, Appl	1669	5	4.2	22	2	US-08-723-425A-55	Sequence 55, Appl
1597	5	4.2	20	2	US-08-802-981-103	Sequence 103, App	1670	5	4.2	22	2	US-08-723-425A-56	Sequence 56, Appl
1598	5	4.2	20	2	US-09-230-421-11	Sequence 11, Appl	1671	5	4.2	22	2	US-08-723-425A-116	Sequence 116, App
1599	5	4.2	20	2	US-09-230-421-12	Sequence 12, Appl	1672	5	4.2	22	2	US-08-723-425A-143	Sequence 143, App
1600	5	4.2	20	2	US-09-411-578-6	Sequence 6, Appl	1673	5	4.2	22	2	US-08-723-425A-143	Sequence 143, App
1601	5	4.2	20	2	US-09-275-265-9	Sequence 9, Appl	1674	5	4.2	22	2	US-08-723-425A-148	Sequence 148, App
1602	5	4.2	20	2	US-09-275-265-12	Sequence 12, Appl	1675	5	4.2	22	2	US-08-723-425A-150	Sequence 150, App
1603	5	4.2	20	2	US-09-275-265-13	Sequence 13, Appl	1676	5	4.2	22	2	US-08-723-425A-152	Sequence 152, App
1604	5	4.2	20	2	US-09-275-265-14	Sequence 14, Appl	1677	5	4.2	22	2	US-09-112-206-49	Sequence 49, Appl
1605	5	4.2	20	2	US-09-275-265-15	Sequence 15, Appl	1678	5	4.2	22	2	US-09-112-206-53	Sequence 53, Appl
1606	5	4.2	20	2	US-08-850-328-9	Sequence 9, Appl	1679	5	4.2	22	2	US-09-112-206-54	Sequence 54, Appl
1607	5	4.2	20	2	US-08-850-328-10	Sequence 10, Appl	1680	5	4.2	22	2	US-09-112-206-55	Sequence 55, Appl
1608	5	4.2	20	2	US-09-941-611-9	Sequence 9, Appl	1681	5	4.2	22	2	US-09-112-206-56	Sequence 56, Appl
1609	5	4.2	20	2	US-09-941-611-12	Sequence 12, Appl	1682	5	4.2	22	2	US-09-112-206-116	Sequence 116, App
1610	5	4.2	20	2	US-09-941-611-13	Sequence 13, Appl	1683	5	4.2	22	2	US-09-112-206-142	Sequence 142, App
1611	5	4.2	20	2	US-09-941-611-14	Sequence 14, Appl	1684	5	4.2	22	2	US-09-112-206-143	Sequence 143, App
1612	5	4.2	20	2	US-09-941-611-15	Sequence 15, Appl	1685	5	4.2	22	2	US-09-112-206-148	Sequence 148, App
1613	5	4.2	20	2	US-09-790-497A-49	Sequence 49, Appl	1686	5	4.2	22	2	US-09-112-206-150	Sequence 150, App
1614	5	4.2	20	2	US-09-790-497A-53	Sequence 53, Appl	1687	5	4.2	22	2	US-09-112-206-152	Sequence 152, App
1615	5	4.2	20	2	US-08-790-497A-54	Sequence 54, Appl	1688	5	4.2	22	2	US-09-576-824A-49	Sequence 49, Appl
1616	5	4.2	20	2	US-09-790-497A-55	Sequence 55, Appl	1689	5	4.2	22	2	US-09-576-824A-53	Sequence 53, Appl
1617	5	4.2	20	2	US-09-790-497A-56	Sequence 56, Appl	1690	5	4.2	22	2	US-09-576-824A-54	Sequence 54, Appl
1618	5	4.2	20	2	US-09-790-497A-116	Sequence 116, App	1691	5	4.2	22	2	US-09-576-824A-55	Sequence 55, Appl
1619	5	4.2	20	2	US-09-790-497A-142	Sequence 142, App	1692	5	4.2	22	2	US-09-576-824A-56	Sequence 56, Appl
1620	5	4.2	20	2	US-09-790-497A-143	Sequence 143, App	1693	5	4.2	22	2	US-09-680-497-49	Sequence 49, Appl
1621	5	4.2	20	2	US-09-790-497A-148	Sequence 148, App	1694	5	4.2	22	2	US-09-680-497-53	Sequence 53, Appl
1622	5	4.2	20	2	US-09-790-497A-150	Sequence 150, App	1695	5	4.2	22	2	US-09-680-497-54	Sequence 54, Appl
1623	5	4.2	20	2	US-09-790-497A-152	Sequence 152, App	1696	5	4.2	22	2	US-09-680-497-55	Sequence 55, Appl
1624	5	4.2	20	2	US-09-576-824A-116	Sequence 116, App	1697	5	4.2	22	2	US-09-680-497-56	Sequence 56, Appl
1625	5	4.2	20	2	US-09-576-824A-142	Sequence 142, App	1698	5	4.2	22	2	US-09-680-497-116	Sequence 116, App
1626	5	4.2	20	2	US-09-576-824A-143	Sequence 143, App	1699	5	4.2	22	2	US-09-680-497-142	Sequence 142, App
1627	5	4.2	20	2	US-09-576-824A-148	Sequence 148, App	1700	5	4.2	22	2	US-09-680-497-143	Sequence 143, App
1628	5	4.2	20	2	US-09-576-824A-150	Sequence 150, App	1701	5	4.2	22	2	US-09-680-497-148	Sequence 148, App
1629	5	4.2	20	2	US-09-576-824A-152	Sequence 152, App	1702	5	4.2	22	2	US-09-680-497-150	Sequence 150, App
1630	5	4.2	20	2	US-09-749-233-6	Sequence 6, Appl	1703	5	4.2	22	2	US-09-680-497-152	Sequence 152, App
1631	5	4.2	20	2	US-09-878-281A-102	Sequence 102, App	1704	5	4.2	23	1	US-08-631-421-7	Sequence 7, Appl
1632	5	4.2	20	2	US-09-878-281A-272	Sequence 272, App	1705	5	4.2	23	2	US-08-731-336-1	Sequence 1, Appl
1633	5	4.2	20	2	US-09-878-281A-273	Sequence 273, App	1706	5	4.2	23	2	US-09-257-667-1	Sequence 1, Appl

1707	5	4.2	23	2	US-09-149-476-328	Sequence 328, App	1780	5	4.2	33	1	US-08-146-028-57	Sequence 57, Appl
1708	5	4.2	23	2	US-09-881-239-8	Sequence 8, Appli	1781	5	4.2	33	1	US-08-146-028-106	Sequence 106, App
1709	5	4.2	23	2	US-09-881-654-7	Sequence 7, Appli	1782	5	4.2	33	1	US-08-146-028-144	Sequence 144, App
1710	5	4.2	23	2	US-10-015-328-6	Sequence 6, Appli	1783	5	4.2	33	2	US-08-723-425A-57	Sequence 57, Appl
1711	5	4.2	23	2	US-09-728-653-6	Sequence 6, Appli	1784	5	4.2	33	2	US-08-723-425A-106	Sequence 106, App
1712	5	4.2	23	2	US-10-637-323-7	Sequence 7, Appli	1785	5	4.2	33	2	US-08-723-425A-144	Sequence 144, App
1713	5	4.2	23	2	US-09-774-639-200	Sequence 200, App	1786	5	4.2	33	2	US-08-604-365-2	Sequence 2, Appli
1714	5	4.2	23	2	US-09-623-548A-1512	Sequence 1512, Ap	1787	5	4.2	33	2	US-09-112-206-57	Sequence 57, Appl
1715	5	4.2	23	2	US-09-657-276-1512	Sequence 1512, Ap	1788	5	4.2	33	2	US-09-112-206-106	Sequence 106, App
1716	5	4.2	23	2	US-08-628-291-14	Sequence 14, Appl	1789	5	4.2	33	2	US-09-112-206-144	Sequence 144, App
1717	5	4.2	24	1	US-08-631-421-6	Sequence 6, Appli	1790	5	4.2	33	2	US-09-586-563C-9	Sequence 9, Appli
1718	5	4.2	24	1	US-09-128-722-14	Sequence 14, Appl	1791	5	4.2	33	2	US-09-586-562C-9	Sequence 2, Appli
1719	5	4.2	24	2	US-08-641-873-11	Sequence 11, Appl	1792	5	4.2	33	2	US-09-689-678-2	Sequence 2, Appli
1720	5	4.2	24	2	US-08-983-075D-11	Sequence 11, Appl	1793	5	4.2	33	2	US-09-790-497A-452	Sequence 452, App
1721	5	4.2	24	2	US-09-270-767-39936	Sequence 39936, A	1794	5	4.2	33	2	US-09-576-824A-57	Sequence 57, Appl
1722	5	4.2	24	2	US-09-270-767-55153	Sequence 55153, A	1795	5	4.2	33	2	US-09-576-824A-109	Sequence 109, App
1723	5	4.2	25	1	US-08-631-421-1	Sequence 1, Appli	1796	5	4.2	33	2	US-09-576-824A-452	Sequence 452, App
1724	5	4.2	25	1	US-08-631-421-4	Sequence 4, Appli	1797	5	4.2	33	2	US-09-680-497-106	Sequence 106, App
1725	5	4.2	25	1	US-08-631-421-5	Sequence 5, Appli	1798	5	4.2	33	2	US-09-680-497-144	Sequence 144, App
1726	5	4.2	25	1	US-08-765-452-15	Sequence 15, Appl	1799	5	4.2	33	2	US-08-190-802A-262	Sequence 262, App
1727	5	4.2	25	2	US-08-641-873-15	Sequence 15, Appl	1800	5	4.2	34	1	US-08-652-450A-2	Sequence 2, Appli
1728	5	4.2	25	2	US-09-049-691-58	Sequence 58, Appl	1801	5	4.2	34	1	US-08-477-346-262	Sequence 262, App
1729	5	4.2	25	2	US-09-341-833A-4	Sequence 4, Appli	1802	5	4.2	34	2	US-08-473-089-262	Sequence 262, App
1730	5	4.2	25	2	US-09-291-925A-9	Sequence 9, Appli	1803	5	4.2	34	2	US-08-487-072A-262	Sequence 262, App
1731	5	4.2	25	2	US-09-988-842-5	Sequence 5, Appli	1804	5	4.2	34	2	US-09-779-451-70	Sequence 70, Appl
1732	5	4.2	25	2	US-09-929-955-39	Sequence 39, Appl	1805	5	4.2	34	2	US-09-576-824A-404	Sequence 404, App
1733	5	4.2	25	2	US-09-930-591-24	Sequence 24, Appl	1806	5	4.2	34	2	US-09-595-682B-1	Sequence 1, Appli
1734	5	4.2	26	1	US-08-232-453A-47	Sequence 47, Appl	1807	5	4.2	34	2	US-09-605-042A-45	Sequence 45, Appl
1735	5	4.2	26	1	US-07-942-245-377	Sequence 377, App	1808	5	4.2	34	2	US-08-337-483-61	Sequence 61, Appl
1736	5	4.2	26	1	US-08-631-421-2	Sequence 2, Appli	1809	5	4.2	35	1	US-08-478-373-61	Sequence 61, Appl
1737	5	4.2	26	1	US-08-631-421-3	Sequence 3, Appli	1810	5	4.2	35	1	US-08-478-373-61	Sequence 61, Appl
1738	5	4.2	26	1	US-08-985-090-7	Sequence 7, Appli	1811	5	4.2	35	1	US-08-487-890A-61	Sequence 61, Appl
1739	5	4.2	26	2	US-09-165-543-7	Sequence 14, Appl	1812	5	4.2	35	1	US-08-652-450A-1	Sequence 1, Appli
1740	5	4.2	26	2	US-09-165-543-14	Sequence 14, Appl	1813	5	4.2	35	1	US-08-652-450A-3	Sequence 3, Appli
1741	5	4.2	26	2	US-09-270-767-59654	Sequence 59654, A	1814	5	4.2	35	1	US-08-750-194-1	Sequence 1, Appli
1742	5	4.2	27	1	US-08-343-427B-1	Sequence 1, Appli	1815	5	4.2	35	1	US-08-478-435-61	Sequence 61, Appl
1743	5	4.2	27	1	US-08-343-427B-2	Sequence 2, Appli	1816	5	4.2	35	1	US-08-337-483-61	Sequence 61, Appl
1744	5	4.2	27	1	US-08-343-427B-3	Sequence 3, Appli	1817	5	4.2	35	1	US-08-478-373-61	Sequence 61, Appl
1745	5	4.2	27	1	US-08-343-427B-4	Sequence 4, Appli	1818	5	4.2	35	2	US-08-478-373-61	Sequence 61, Appl
1746	5	4.2	27	1	US-08-343-427B-5	Sequence 5, Appli	1819	5	4.2	35	2	US-08-478-373-61	Sequence 61, Appl
1747	5	4.2	27	1	US-08-343-427B-6	Sequence 6, Appli	1820	5	4.2	35	2	US-08-483-577A-61	Sequence 61, Appl
1748	5	4.2	27	1	US-08-343-427B-7	Sequence 7, Appli	1821	5	4.2	35	2	US-08-897-438-61	Sequence 61, Appl
1749	5	4.2	27	1	US-08-343-427B-8	Sequence 8, Appli	1822	5	4.2	35	2	US-08-637-654-61	Sequence 61, Appl
1750	5	4.2	27	1	US-08-343-427B-9	Sequence 9, Appli	1823	5	4.2	35	2	US-08-649-518-61	Sequence 61, Appl
1751	5	4.2	27	1	US-08-343-427B-10	Sequence 10, Appl	1824	5	4.2	35	2	US-09-270-767-58973	Sequence 58973, A
1752	5	4.2	27	1	US-08-631-421-8	Sequence 8, Appli	1825	5	4.2	35	6	US-09-788-308E-3	Sequence 3, Appli
1753	5	4.2	27	1	US-08-283-917-11	Sequence 11, Appl	1826	5	4.2	35	6	US-09-605-042A-45	Sequence 45, Appl
1754	5	4.2	27	1	US-08-961-716-11	Sequence 11, Appl	1827	5	4.2	35	6	US-08-487-890A-61	Sequence 61, Appl
1755	5	4.2	27	1	US-08-961-716-11	Sequence 11, Appl	1828	5	4.2	35	6	US-08-487-890A-61	Sequence 61, Appl
1756	5	4.2	27	2	US-09-053-197A-47	Sequence 47, Appl	1829	5	4.2	35	6	US-08-652-450A-1	Sequence 1, Appli
1757	5	4.2	27	2	US-09-116-294-3	Sequence 3, Appli	1830	5	4.2	35	6	US-08-652-450A-3	Sequence 3, Appli
1758	5	4.2	27	2	US-09-085-761A-52	Sequence 52, Appl	1831	5	4.2	36	1	US-08-750-194-1	Sequence 1, Appli
1759	5	4.2	27	2	US-09-085-761A-52	Sequence 52, Appl	1832	5	4.2	36	1	US-08-478-435-61	Sequence 61, Appl
1760	5	4.2	27	2	US-09-085-761A-52	Sequence 52, Appl	1833	5	4.2	36	2	US-08-478-435-61	Sequence 61, Appl
1761	5	4.2	27	2	US-09-512-260A-4	Sequence 20, Appl	1834	5	4.2	37	1	US-07-649-591B-1	Sequence 1, Appli
1762	5	4.2	28	2	US-10-144-549-16	Sequence 16, Appl	1835	5	4.2	37	1	US-08-277-540-1	Sequence 1, Appli
1763	5	4.2	28	2	US-08-995-227A-1	Sequence 1, Appli	1836	5	4.2	37	1	US-08-430-787A-1	Sequence 1, Appli
1764	5	4.2	28	2	US-09-019-095A-34	Sequence 34, Appl	1837	5	4.2	38	1	US-08-488-161-3	Sequence 3, Appli
1765	5	4.2	28	2	US-09-020-846-42	Sequence 42, Appl	1838	5	4.2	38	2	US-09-273-685-3	Sequence 3, Appli
1766	5	4.2	28	2	US-10-049-113A-2	Sequence 2, Appli	1839	5	4.2	38	4	PCT-US95-11934-3	Sequence 3, Appli
1767	5	4.2	29	2	US-08-641-873-10	Sequence 10, Appl	1840	5	4.2	39	2	US-09-188-930-133	Sequence 133, App
1768	5	4.2	30	6	5188961-8	Patent No. 5188961	1841	5	4.2	39	2	US-09-220-528-48	Sequence 48, Appl
1769	5	4.2	31	2	US-08-466-368-17	Sequence 17, Appl	1842	5	4.2	39	2	US-09-020-846-41	Sequence 41, Appl
1770	5	4.2	31	2	US-08-470-998-14	Sequence 14, Appl	1843	5	4.2	40	2	US-09-312-283C-133	Sequence 133, App
1771	5	4.2	31	2	US-08-776-188C-66	Sequence 66, Appl	1844	5	4.2	40	2	US-09-037-524-4	Sequence 4, Appli
1772	5	4.2	31	2	US-08-790-437A-57	Sequence 57, Appl	1845	5	4.2	40	2	US-09-270-767-39315	Sequence 39315, A
1773	5	4.2	31	2	US-08-790-497A-109	Sequence 109, App	1846	5	4.2	40	2	US-09-270-767-54532	Sequence 54532, A
1774	5	4.2	31	2	US-08-328-500-22	Sequence 22, Appl	1847	5	4.2	40	2	US-09-385-219A-67	Sequence 67, Appl
1775	5	4.2	31	2	US-09-270-767-37053	Sequence 37053, A	1848	5	4.2	40	2	US-09-392-476-4	Sequence 4, Appli
1776	5	4.2	32	1	US-08-652-450A-4	Sequence 4, Appli	1849	5	4.2	41	2	US-09-270-767-58649	Sequence 58649, A
1777	5	4.2	32	2	US-08-790-437A-144	Sequence 144, App	1850	5	4.2	42	1	US-08-441-871-82	Sequence 82, Appli
1778	5	4.2	32	2	US-08-790-497A-404	Sequence 404, App	1851	5	4.2	43	1	US-08-162-149-6	Sequence 6, Appli
1779	5	4.2	32	2	US-09-576-824A-144	Sequence 144, App	1852	5	4.2	43	2	US-09-205-258-296	Sequence 296, App
												US-09-270-767-60200	Sequence 60200, A

1853	5	4.2	43	2	US-10-172-502-23	Sequence 23, Appl	1926	5	4.2	60	2	US-09-621-976-6703	Sequence 6703, Ap
1854	5	4.2	43	2	US-10-004-860-296	Sequence 296, App	1927	5	4.2	60	2	US-09-583-110-5122	Sequence 5122, Ap
1855	5	4.2	44	2	US-09-491-614B-35	Sequence 35, Appl	1928	5	4.2	60	2	US-09-248-796A-27981	Sequence 27981, A
1856	5	4.2	44	6	5188961-4	Patent No. 5188961	1929	5	4.2	60	2	US-09-513-999C-6335	Sequence 6335, Ap
1857	5	4.2	45	1	US-07-689-693B-19	Sequence 19, Appl	1930	5	4.2	60	2	US-09-471-276-869	Sequence 869, App
1858	5	4.2	45	1	US-07-689-693B-20	Sequence 20, Appl	1931	5	4.2	60	2	US-09-902-540-11781	Sequence 11781, A
1859	5	4.2	45	1	US-08-530-290-22	Sequence 22, Appl	1932	5	4.2	60	2	US-09-902-540-14615	Sequence 14615, A
1860	5	4.2	45	1	US-09-461-325-163	Sequence 163, App	1933	5	4.2	60	6	5188961-1	Patent No. 5188961
1861	5	4.2	45	2	US-10-012-425-163	Sequence 163, App	1934	5	4.2	61	2	US-09-540-236-2368	Sequence 2368, Ap
1862	5	4.2	45	2	US-09-270-767-37178	Sequence 37178, A	1935	5	4.2	61	2	US-09-270-767-57597	Sequence 57597, A
1863	5	4.2	45	2	US-09-270-767-52395	Sequence 52395, A	1936	5	4.2	62	1	US-08-381-041A-1	Sequence 1, Appl
1864	5	4.2	45	2	US-09-270-767-52395	Sequence 52395, A	1937	5	4.2	62	2	US-08-630-915A-135	Sequence 135, App
1865	5	4.2	45	2	US-09-270-767-59947	Sequence 59947, A	1938	5	4.2	62	2	US-09-621-976-4501	Sequence 4501, Ap
1866	5	4.2	45	2	US-10-115-123-163	Sequence 163, App	1939	5	4.2	62	2	US-09-270-767-61619	Sequence 61619, A
1867	5	4.2	46	1	US-08-383-631-2	Sequence 2, Appl	1940	5	4.2	62	2	US-09-879-957-135	Sequence 135, App
1868	5	4.2	46	1	US-08-178-477B-20	Sequence 20, Appl	1941	5	4.2	62	2	US-09-513-999C-5977	Sequence 5977, Ap
1869	5	4.2	46	2	US-09-779-451-69	Sequence 69, Appl	1942	5	4.2	62	2	US-09-513-999C-5979	Sequence 5979, Ap
1870	5	4.2	46	2	US-09-270-767-32839	Sequence 32839, A	1943	5	4.2	63	1	US-08-381-049-1	Sequence 1, Appl
1871	5	4.2	46	2	US-09-270-767-48056	Sequence 48056, A	1944	5	4.2	63	1	US-08-644-664B-34	Sequence 34, Appl
1872	5	4.2	47	2	US-09-205-258-443	Sequence 443, App	1945	5	4.2	63	1	US-08-761-277A-34	Sequence 34, Appl
1873	5	4.2	47	2	US-10-004-860-443	Sequence 443, App	1946	5	4.2	63	2	US-09-443-041A-14	Sequence 14, Appl
1874	5	4.2	48	1	US-07-796-361A-15	Sequence 15, Appl	1947	5	4.2	63	2	US-09-248-796A-23058	Sequence 23058, A
1875	5	4.2	48	2	US-08-804-439A-105	Sequence 105, App	1948	5	4.2	64	2	US-09-489-039A-8930	Sequence 8930, A
1876	5	4.2	49	1	US-08-398-021-2	Sequence 2, Appl	1949	5	4.2	64	2	US-09-248-796A-24546	Sequence 24546, A
1877	5	4.2	50	1	US-08-133-011-17	Sequence 17, Appl	1950	5	4.2	64	2	US-09-248-796A-27240	Sequence 27240, A
1878	5	4.2	50	1	US-08-322-730A-17	Sequence 17, Appl	1951	5	4.2	64	2	US-09-248-796A-27368	Sequence 27368, A
1879	5	4.2	50	1	US-08-387-874-17	Sequence 17, Appl	1952	5	4.2	64	2	US-09-902-540-16124	Sequence 16124, A
1880	5	4.2	50	1	US-08-383-619-17	Sequence 17, Appl	1953	5	4.2	65	2	US-08-905-223-403	Sequence 403, App
1881	5	4.2	50	2	US-08-907-739-17	Sequence 17, Appl	1954	5	4.2	65	2	US-08-918-148-73	Sequence 73, Appl
1882	5	4.2	50	2	US-09-729-597-17	Sequence 17, Appl	1955	5	4.2	65	2	US-09-254-776B-78	Sequence 78, Appl
1883	5	4.2	50	2	US-08-843-076D-42	Sequence 42, Appl	1956	5	4.2	65	2	US-09-107-433-3854	Sequence 3854, Ap
1884	5	4.2	50	2	US-09-621-976-4324	Sequence 4324, Ap	1957	5	4.2	65	2	US-09-107-532A-6494	Sequence 6494, Ap
1885	5	4.2	51	4	PCT-US93-08364-17	Sequence 17, Appl	1958	5	4.2	65	2	US-09-270-767-59016	Sequence 59016, A
1886	5	4.2	51	2	US-08-995-227A-2	Sequence 2, Appl	1959	5	4.2	65	2	US-09-248-796A-21245	Sequence 21245, A
1887	5	4.2	51	2	US-09-513-999C-6505	Sequence 6505, Ap	1960	5	4.2	65	2	US-09-513-999C-4565	Sequence 4565, Ap
1888	5	4.2	51	2	US-09-513-999C-7145	Sequence 7145, Ap	1961	5	4.2	65	2	US-09-513-999C-6441	Sequence 6441, Ap
1889	5	4.2	51	2	US-09-902-540-10095	Sequence 10095, A	1962	5	4.2	66	2	US-09-107-433-3854	Sequence 3854, Ap
1890	5	4.2	52	1	US-09-902-540-13154	Sequence 13154, A	1963	5	4.2	66	2	US-09-138-091A-71	Sequence 71, Appl
1891	5	4.2	52	1	US-08-361-920-17	Sequence 17, Appl	1964	5	4.2	66	2	US-09-248-796A-25848	Sequence 25848, A
1892	5	4.2	52	1	US-08-381-666-1	Sequence 1, Appl	1965	5	4.2	67	2	US-09-902-540-16512	Sequence 16512, A
1893	5	4.2	52	1	US-08-479-939-17	Sequence 17, Appl	1966	5	4.2	67	2	US-09-107-433-2792	Sequence 2792, Ap
1894	5	4.2	52	1	US-08-483-432-17	Sequence 17, Appl	1967	5	4.2	68	2	US-09-902-540-12697	Sequence 12697, A
1895	5	4.2	53	2	US-09-205-258-983	Sequence 983, App	1968	5	4.2	69	2	US-09-489-039A-8424	Sequence 8424, A
1896	5	4.2	53	2	US-09-621-976-6224	Sequence 6224, Ap	1969	5	4.2	69	2	US-08-858-207A-364	Sequence 364, App
1897	5	4.2	54	1	US-10-004-860-983	Sequence 983, App	1970	5	4.2	69	2	US-09-252-991A-24397	Sequence 24397, A
1898	5	4.2	54	1	US-08-381-370-1	Sequence 1, Appl	1971	5	4.2	69	2	US-09-513-999C-4505	Sequence 4505, Ap
1899	5	4.2	54	2	US-09-621-976-6167	Sequence 6167, Ap	1972	5	4.2	70	1	US-08-162-149-4	Sequence 4, Appl
1900	5	4.2	54	2	US-09-621-976-6203	Sequence 6203, Ap	1973	5	4.2	70	2	US-09-134-000C-3856	Sequence 3856, Ap
1901	5	4.2	54	2	US-09-270-767-33403	Sequence 33403, A	1974	5	4.2	70	2	US-09-621-976-5545	Sequence 5545, Ap
1902	5	4.2	54	2	US-08-270-767-8620	Sequence 8620, A	1975	5	4.2	70	2	US-09-621-976-5648	Sequence 5648, Ap
1903	5	4.2	54	2	US-09-513-999C-5831	Sequence 5831, App	1976	5	4.2	70	2	US-09-300, App	Sequence 300, App
1904	5	4.2	54	2	US-09-595-682B-3	Sequence 3, Appl	1977	5	4.2	70	2	US-09-289-346B-3	Sequence 3, Appl
1905	5	4.2	54	2	US-09-595-682B-5	Sequence 5, Appl	1978	5	4.2	70	2	US-09-289-346B-14	Sequence 14, Appl
1906	5	4.2	55	1	US-09-562-930-8	Sequence 8, Appl	1979	5	4.2	71	1	US-07-689-693B-1	Sequence 1, Appl
1907	5	4.2	55	1	US-08-102-385G-14	Sequence 14, Appl	1980	5	4.2	71	2	US-09-056-556-233	Sequence 233, App
1908	5	4.2	55	2	US-09-621-976-6675	Sequence 6675, Ap	1981	5	4.2	71	2	US-09-072-596-228	Sequence 228, App
1909	5	4.2	55	2	US-09-513-999C-5978	Sequence 5978, Ap	1982	5	4.2	71	2	US-09-252-991A-30306	Sequence 30306, A
1910	5	4.2	55	2	US-09-513-999C-5980	Sequence 5980, Ap	1983	5	4.2	71	2	US-09-072-596-233	Sequence 233, App
1911	5	4.2	55	2	US-09-513-999C-5981	Sequence 5981, Ap	1984	5	4.2	71	2	US-09-489-039A-12114	Sequence 12114, A
1912	5	4.2	56	2	US-09-471-276-1527	Sequence 1527, Ap	1985	5	4.2	71	2	US-09-621-976-7588	Sequence 7588, Ap
1913	5	4.2	56	2	US-09-513-999C-6755	Sequence 6755, Ap	1986	5	4.2	71	2	US-09-621-976-7596	Sequence 7596, Ap
1914	5	4.2	56	2	US-09-513-999C-7036	Sequence 7036, Ap	1987	5	4.2	71	2	US-09-621-976-7612	Sequence 7612, Ap
1915	5	4.2	56	2	US-09-902-540-14481	Sequence 14481, A	1988	5	4.2	71	2	US-09-621-976-7624	Sequence 7624, Ap
1916	5	4.2	57	2	US-09-621-976-7041	Sequence 7041, App	1989	5	4.2	71	2	US-09-621-976-7664	Sequence 7664, Ap
1917	5	4.2	58	2	US-09-461-325-166	Sequence 166, App	1990	5	4.2	71	2	US-09-621-976-7683	Sequence 7683, Ap
1918	5	4.2	58	2	US-10-012-542-166	Sequence 166, App	1991	5	4.2	71	2	US-09-621-976-7687	Sequence 7687, Ap
1919	5	4.2	58	2	US-09-270-767-41119	Sequence 41119, A	1992	5	4.2	71	2	US-09-513-999C-4570	Sequence 4570, Ap
1920	5	4.2	58	2	US-09-270-767-56335	Sequence 56335, A	1993	5	4.2	71	2	US-09-513-999C-7498	Sequence 7498, Ap
1921	5	4.2	58	2	US-10-113-123-166	Sequence 166, App	1994	5	4.2	71	2	US-10-193-002-228	Sequence 228, App
1922	5	4.2	58	2	US-09-513-999C-4523	Sequence 4523, Ap	1995	5	4.2	71	2	US-10-084-843-233	Sequence 233, App
1923	5	4.2	59	2	US-10-001-887-124	Sequence 124, App	1996	5	4.2	72	2	US-08-995-227A-3	Sequence 3, Appl
1924	5	4.2	59	2	US-09-621-976-6080	Sequence 6080, Ap	1997	5	4.2	72	2	US-09-270-767-59575	Sequence 59575, A
1925	5	4.2	59	2	US-09-583-110-4468	Sequence 4468, Ap	1998	5	4.2	72	2	US-09-866-073A-17	Sequence 17, Appl

1999 5 4.2 72 2 US-09-513-999C-7097 Sequence 7097, Ap
2000 5 4.2 73 1 US-07-689-693B-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-878-281A-36
; Sequence 36, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 36
; LENGTH: 133
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-36

Query Match 100.0%; Score 118; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 4.9e-106;
Matches 118; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 60
Db 16 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 75

Qy 61 E M E C S Q A A P Y I E Q A Q V I A H Q F K G K V L G L L Q R A T O Q Q A V I P I V T T N W K L E A F W H K H 118
Db 76 E M E C S Q A A P Y I E Q A Q V I A H Q F K G K V L G L L Q R A T O Q Q A V I P I V T T N W K L E A F W H K H 133

RESULT 2
US-09-878-281A-38
; Sequence 38, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
; LENGTH: 133
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-38

Query Match 70.3%; Score 83; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.5e-72;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 60
Db 16 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 75

Qy 61 E M E C S Q A A P Y I E Q A Q V I A H Q F K 83
Db 76 E M E C S Q A A P Y I E Q A Q V I A H Q F K 98

RESULT 3
US-09-878-281A-40

; Sequence 40, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 133
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-40

Query Match 70.3%; Score 83; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 2.5e-72;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 60
Db 16 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 75

Qy 61 E M E C S Q A A P Y I E Q A Q V I A H Q F K 83
Db 76 E M E C S Q A A P Y I E Q A Q V I A H Q F K 98

RESULT 4
US-09-878-281A-223
; Sequence 223, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 223
; LENGTH: 209
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-223

Query Match 70.3%; Score 83; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 3.8e-72;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 60
Db 92 A C M S A D L E V T T S T W L L G G V L A A A Y C L S V G C V V I V G H I E L G G K P A I V P D K E V L Y Q Q Y D 151

Qy 61 E M E C S Q A A P Y I E Q A Q V I A H Q F K 83
Db 152 E M E C S Q A A P Y I E Q A Q V I A H Q F K 174

RESULT 5
US-08-244-116B-17
; Sequence 17, Application US/08244116B
; Patent No. 5763159
; GENERAL INFORMATION:
; APPLICANT: Simmonds, Peter
; APPLICANT: Chan, Shiu-Wan
; APPLICANT: Yap, Peng L.
; TITLE OF INVENTION: Hepatitis-C Virus Testing
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell, Seltzer, Park & Gibson, P.A.
STREET: 1211 East Morehead Street
CITY: Charlotte
STATS: No. 5763159th Carolina
COUNTRY: United States
ZIP: 28234
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0. Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,116B
FILING DATE: 15-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB92/02143
FILING DATE: 20-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Sibley, Kenneth D.
REGISTRATION NUMBER: 31,665
REFERENCE/DOCKET NUMBER: 1749-125
TELECOMMUNICATION INFORMATION:
TELEPHONE: 704-377-1561
TELEFAX: 704-334-2014
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 128 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: yes
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Hepatitis C virus
US-08-244-116B-17

Query Match 39.8%; Score 47; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 1.2e-37;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 47
|||||
Db 10 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 56
|||||

RESULT 6
US-09-878-281A-32
Sequence 32, Application US/09878281A
Patent No. 6762024
GENERAL INFORMATION:
APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
FILE REFERENCE: 35
CURRENT APPLICATION NUMBER: US/09/878,281A
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 32
LENGTH: 133
TYPE: PRT
ORGANISM: hepatitis C virus
US-09-878-281A-32

Query Match 39.8%; Score 47; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.2e-37;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 47
|||||
Db 16 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 62
|||||

RESULT 7
US-09-878-281A-34
Sequence 34, Application US/09878281A
Patent No. 6762024
GENERAL INFORMATION:
APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
FILE REFERENCE: 35
CURRENT APPLICATION NUMBER: US/09/878,281A
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34
LENGTH: 133
TYPE: PRT
ORGANISM: hepatitis C virus
US-09-878-281A-34

Query Match 39.8%; Score 47; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.2e-37;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 47
|||||
Db 16 ACMSADLEVTSTWVLGGVLAALAAAYCLSVGCVVIVGHIELGKPA 62
|||||

RESULT 8
US-09-881-239-5
Sequence 5, Application US/09881239
Patent No. 6630298
GENERAL INFORMATION:
APPLICANT: CHIEN, David Y.
APPLICANT: TANDESKE, Laura
APPLICANT: GEORGE-NASCIEMENTO, Carlos
APPLICANT: COIT, Doris
APPLICANT: MEDINA-SELBY, Angelica
TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
FILE REFERENCE: 2302-16073 / PFI6073.003
CURRENT APPLICATION NUMBER: US/09/881,239
CURRENT FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 829
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: MEFA 12
US-09-881-239-5

Query Match 34.7%; Score 41; DB 2; Length 829;
Best Local Similarity 100.0%; Pred. No. 3.9e-31;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GGGPAIVDPKVELVYQQYDEMEBCSQAPYIEQAQVIAHQFK 83
|||||
Db 444 GGGPAIVDPKVELVYQQYDEMEBCSQAPYIEQAQVIAHQFK 484
|||||

RESULT 9
US-09-881-654-4
Sequence 4, Application US/09881654
Patent No. 6632601
GENERAL INFORMATION:
APPLICANT: CHIEN, David Y.
APPLICANT: TANDESKE, Laura
APPLICANT: GEORGE-NASCIEMENTO, Carlos
US-09-881-654-4

APPLICANT: COLT, Doris
APPLICANT: MEDINA-SELBY, Angelica
TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
FILE REFERENCE: 2302-17039 / PPI7039.002
CURRENT APPLICATION NUMBER: US/09/881,654
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 60/212,082
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/280,811
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/280,867
PRIOR FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1099
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1
US-09-881-654-4

Query Match 34.7%; Score 41; DB 2; Length 1099;
Best Local Similarity 100.0%; Pred. No. 5.1e-31;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GGKPAIVDPKEVLYQQYDEMECSQAAPYIEQAQVIAHQFK 83
|||||
Db 748 GGKPAIVDPKEVLYQQYDEMECSQAAPYIEQAQVIAHQFK 788

RESULT 10
US-10-637-323-4
Sequence 4, Application US/10637323
Patent No. 6797809
GENERAL INFORMATION:
APPLICANT: CHIEN, David Y.
APPLICANT: ARANGEL, Philip
APPLICANT: TANDESKE, Laura
APPLICANT: GEORGE-NASCIEMENTO, Carlos
APPLICANT: COLT, Doris
APPLICANT: MEDINA-SELBY, Angelica
TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
FILE REFERENCE: 2302-17039 / PPI7039.002
CURRENT APPLICATION NUMBER: US/10/637,323
CURRENT FILING DATE: 2003-08-08
PRIOR APPLICATION NUMBER: US/09/881,654
PRIOR FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: 60/212,082
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 60/280,811
PRIOR FILING DATE: 2001-04-02
PRIOR APPLICATION NUMBER: 60/280,867
PRIOR FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 1099
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: MEFA 7.1
US-10-637-323-4

Query Match 34.7%; Score 41; DB 2; Length 1099;
Best Local Similarity 100.0%; Pred. No. 5.1e-31;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GGKPAIVDPKEVLYQQYDEMECSQAAPYIEQAQVIAHQFK 83
|||||
Db 748 GGKPAIVDPKEVLYQQYDEMECSQAAPYIEQAQVIAHQFK 788

RESULT 11

US-09-790-497A-104
Sequence 104, Application US/09790497A
Patent No. 6649735
GENERAL INFORMATION:
APPLICANT: De Leys, Robert
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
TITLE OF INVENTION: CONTAINING THEM
FILE REFERENCE: 2752-16
CURRENT APPLICATION NUMBER: US/09/790,497A
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 09/576,824
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 08/723,425
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: 09/146,028
PRIOR FILING DATE: 1993-11-22
PRIOR APPLICATION NUMBER: PCT/EP93/00517
PRIOR FILING DATE: 1993-03-08
PRIOR APPLICATION NUMBER: EP 92400598.6
PRIOR FILING DATE: 1992-03-06
NUMBER OF SEQ ID NOS: 600
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 104
LENGTH: 20
TYPE: PRT
ORGANISM: Hepatitis C virus
US-09-790-497A-104

Query Match 16.9%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVDPKEVLYQQYDE 61
|||||
Db 1 LGGKPAIVDPKEVLYQQYDE 20

RESULT 12

US-09-878-281A-97
Sequence 97, Application US/09878281A
Patent No. 6762024
GENERAL INFORMATION:
APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
TITLE OF INVENTION: and therapy
FILE REFERENCE: 35
CURRENT APPLICATION NUMBER: US/09/878,281A
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 97
LENGTH: 20
TYPE: PRT
ORGANISM: hepatitis C virus
US-09-878-281A-97

Query Match 16.9%; Score 20; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.1e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVDPKEVLYQQYDE 61
|||||
Db 1 LGGKPAIVDPKEVLYQQYDE 20

RESULT 13

US-08-146-028-107
Sequence 107, Application US/08146028

; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-107

Query Match 16.9%; Score 20; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVDPKVELYQQYDE 61
Db 2 LGGKPAIVDPKVELYQQYDE 21

RESULT 14
US-08-723-425A-107
; Sequence 107, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-107

Query Match 16.9%; Score 20; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVDPKVELYQQYDE 61
Db 2 LGGKPAIVDPKVELYQQYDE 21

RESULT 15
US-09-112-206-107
; Sequence 107, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-107

Query Match 16.9%; Score 20; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVDPKVELYQQYDE 61
Db 2 LGGKPAIVDPKVELYQQYDE 21

RESULT 16
US-09-576-824A-104
; Sequence 104, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22

;; PRIOR APPLICATION NUMBER: PCT/EP93/00517
;; PRIOR FILING DATE: 1993-03-08
;; PRIOR APPLICATION NUMBER: EP 92400598.6
;; PRIOR FILING DATE: 1992-03-06
;; NUMBER OF SEQ ID NOS: 600
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 104
;; LENGTH: 22
;; TYPE: PRT
;; ORGANISM: Hepatitis C virus
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (1)
;; OTHER INFORMATION: modified site
;; NAME/KEY: VARIANT
;; LOCATION: (22)
;; OTHER INFORMATION: modified site
US-09-576-824A-104

Query Match 16.9%; Score 20; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVPDKEVLVQYQDE 61
| | | | | | | | | | | | | | | | | | | | | |
Db 2 LGGKPAIVPDKEVLVQYQDE 21

RESULT 17

US-09-680-497-107
; Sequence 107, Application US/09680497
; Patent No. 6709828

; GENERAL INFORMATION:

;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993

; INFORMATION FOR SEQ ID NO: 107:

; SEQUENCE CHARACTERISTICS:

;; LENGTH: 22 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

US-09-680-497-107

Query Match 16.9%; Score 20; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2.3e-12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAIVPDKEVLVQYQDE 61
| | | | | | | | | | | | | | | | | | | | | |
Db 2 LGGKPAIVPDKEVLVQYQDE 21

RESULT 18

US-08-244-116B-2
; Sequence 2, Application US/08244116B
; Patent No. 5763159

;; GENERAL INFORMATION:
;; APPLICANT: Simmonds, Peter
;; APPLICANT: Chan, Shiu-Wan
;; APPLICANT: Yap, Peng L.
;; TITLE OF INVENTION: Hepatitis-C Virus Testing
;; NUMBER OF SEQUENCES: 53
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Bell, Seltzer, Park & Gibson, P.A.
;; STREET: 1211 East Morehead Street
;; CITY: Charlotte
;; STATE: No. 5763159th Carolina
;; COUNTRY: United States
;; ZIP: 28234

; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/244,116B
;; FILING DATE: 15-JUL-1994
;; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/GB92/02143

;; FILING DATE: 20-NOV-1992

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Sibley, Kenneth D.

;; REGISTRATION NUMBER: 31,665

;; REFERENCE/DOCKET NUMBER: 1749-125

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 704-377-1561

;; TELEFAX: 704-334-2014

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 19 amino acids

;; TYPE: amino acid

;; STRANDEDNESS:

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

;; FRAGMENT TYPE: internal

;; ORIGINAL SOURCE:

;; ORGANISM: Hepatitis-C virus

US-08-244-116B-2

Query Match 16.1%; Score 19; DB 1; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 ECSQAAPYIEQAQVIAHQF 82

| | | | | | | | | | | | | | | | | | | | | |
Db 1 ECSQAAPYIEQAQVIAHQF 19

RESULT 19

US-08-537-802-44

;; Sequence 44, Application US/08537802

;; Patent No. 6881821

; GENERAL INFORMATION:

;; APPLICANT:

;; TITLE OF INVENTION: HEPATITIS-C VIRUS TYPE 4, 5 & 6

;; NUMBER OF SEQUENCES: 50

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/537,802

;; FILING DATE:

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/GB94/00957

;; FILING DATE:

; APPLICATION NUMBER: GB 9309237.7
; FILING DATE: 05-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9400263.1
; FILING DATE: 07-JAN-1994
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-537-802-44

Query Match 16.1%; Score 19; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 64 ECSQAAPYIEQAQVIAHQF 82
Db 1 ECSQAAPYIEQAQVIAHQF 19

RESULT 20
US-09-497A-105
; Sequence 105, Application US/09790497A
; Patent No. 6649735

; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A

; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 105
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-497A-105

Query Match 15.3%; Score 18; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SQAAPYIEQAQVIAHQF 83
Db 1 SQAAPYIEQAQVIAHQF 18

RESULT 21
US-09-878-281A-99
; Sequence 99, Application US/09878281A
; Patent No. 6762024

; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35

; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-99

Query Match 15.3%; Score 18; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SQAAPYIEQAQVIAHQF 83
Db 1 SQAAPYIEQAQVIAHQF 18

RESULT 22
US-08-146-028-108
; Sequence 108, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-108

Query Match 15.3%; Score 18; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SQAAPYIEQAQVIAHQF 83
Db 2 SQAAPYIEQAQVIAHQF 19

RESULT 23
US-08-723-425A-108
; Sequence 108, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT

; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA

;; ZIP: 22201
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/723,425A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-13
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-816-4000
;; TELEFAX: 703-816-4100
;; INFORMATION FOR SEQ ID NO: 108:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: Peptide
US-08-723-425A-108

Query Match 15.3%; Score 18; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SOAAPYIEQAQVIAHQFK 83
Db 2 SOAAPYIEQAQVIAHQFK 19

RESULT 24
US-09-112-206-108
; Sequence 108, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-112-206-108

Query Match 15.3%; Score 18; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SOAAPYIEQAQVIAHQFK 83
Db 2 SOAAPYIEQAQVIAHQFK 19

Db 2 SOAAPYIEQAQVIAHQFK 19

RESULT 25

US-09-576-824A-105
; Sequence 105, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 105
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: modified site
; NAME/KEY: VARIANT
; LOCATION: (22)
; OTHER INFORMATION: modified site
US-09-576-824A-105

Query Match 15.3%; Score 18; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SOAAPYIEQAQVIAHQFK 83
Db 2 SOAAPYIEQAQVIAHQFK 19

RESULT 26

US-09-680-497-108
; Sequence 108, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993

Query Match 15.3%; Score 18; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 66 SOAAPYIEQAQVIAHQFK 83
Db 2 SOAAPYIEQAQVIAHQFK 19

```
; INFORMATION FOR SEQ ID NO: 108:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-108

Query Match      15.3%; Score 18; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 SQAAPYIEQAQVIAHQFK 83
Db      2 SQAAPYIEQAQVIAHQFK 19

RESULT 27
US-08-921-887-52
; Sequence 52, Application US/08921887
; Patent No. 6030771
; GENERAL INFORMATION:
; APPLICANT: KHUDYAKOV, YURI E.
; APPLICANT: FIELDS, HOWARD A.
; TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION
; TITLE OF INVENTION: ENDONUCLEASE ASSISTED LIGATION METHOD FOR MAKING THE SAME
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JONES & ASKEW, LLP
; STREET: 191 Peachtree Street, N.W., 37th Floor
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30303-1769
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921.887
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WARREN, WILLIAM L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 03063-0380
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 352 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis virus
US-08-921-887-52

Query Match      15.3%; Score 18; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 SQAAPYIEQAQVIAHQFK 83
Db      125 SQAAPYIEQAQVIAHQFK 142
```

```
RESULT 28
US-09-491-146A-52
; Sequence 52, Application US/09491146A
; Patent No. 6960659
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by
; APPLICANT: the Secretary, Department of Health and Human Services,
; APPLICANT: C/O Centers for Disease Control and Prevention
; APPLICANT: Khudyakov, Yuri
; APPLICANT: Fields, Howard
; TITLE OF INVENTION: MOSAIC PROTEIN AND RESTRICTION ENDONUCLEASE ASSISTED
; TITLE OF INVENTION: LIGATION METHOD FOR MAKING THE SAME
; FILE REFERENCE: 14114.0344U2
; CURRENT APPLICATION NUMBER: US/09/491.146A
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 08/921,887
; PRIOR FILING DATE: 25-08-97
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-491-146A-52

Query Match      15.3%; Score 18; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      66 SQAAPYIEQAQVIAHQFK 83
Db      125 SQAAPYIEQAQVIAHQFK 142

RESULT 29
US-08-244-116B-1
; Sequence 1, Application US/08244116B
; Patent No. 5763159
; GENERAL INFORMATION:
; APPLICANT: Simmonds, Peter
; APPLICANT: Chan, Shiu-Wan
; APPLICANT: Yap, Peng L.
; TITLE OF INVENTION: Hepatitis-C Virus Testing
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5763159th Carolina
; COUNTRY: United States
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244.116B
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02143
; FILING DATE: 20-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 1749-125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 704-377-1561
; TELEFAX: 704-334-2014
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

;
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis-C virus
US-08-244-116B-1

Query Match 11.9%; Score 14; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEM 62
| | | | | | | | | |
Db 5 VPDKEVLYQQYDEM 18

RESULT 30

US-08-537-802-30
; Sequence 30, Application US/08537802
; Patent No. 6881821

; GENERAL INFORMATION:

; APPLICANT: HEPATITIS-C VIRUS TYPE 4, 5 & 6
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; NUMBER OF SEQUENCES: 50
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,802
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/00957
; FILING DATE:
; APPLICATION NUMBER: GB 9309237.7
; FILING DATE: 05-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9400263.1
; FILING DATE: 07-JAN-1994
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-537-802-30

Query Match 11.9%; Score 14; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEM 62
| | | | | | | | | |
Db 5 VPDKEVLYQQYDEM 18

RESULT 31

US-09-878-281A-98
; Sequence 98, Application US/09878281A
; Patent No. 6762024

; GENERAL INFORMATION:

; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/878,281A
; FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-98

Query Match 11.0%; Score 13; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDE 61
| | | | | | | | | |
Db 8 VPDKEVLYQQYDE 20

RESULT 32

US-09-878-281A-100
; Sequence 100, Application US/09878281A
; Patent No. 6762024

; GENERAL INFORMATION:

; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 100
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-100

Query Match 11.0%; Score 13; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLLQRTQQQ 97
| | | | | | | | | |
Db 8 KVLGLLQRTQQQ 20

RESULT 33

US-08-723-425A-109
; Sequence 109, Application US/08723425A
; Patent No. 6165730

; GENERAL INFORMATION:

; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.

; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-109

Query Match 11.0%; Score 13; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQRTQQQ 97
Db 9 KVLGLLQRTQQQ 21

RESULT 34

US-09-790-497A-106
; Sequence 106, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:

; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM

; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06

; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 22
; TYPE: PRT

; ORGANISM: Hepatitis C virus
US-09-790-497A-106

Query Match 11.0%; Score 13; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQRTQQQ 97
Db 10 KVLGLLQRTQQQ 22

RESULT 35

US-08-146-028-109
; Sequence 109, Application US/08146028
; Patent No. 5891840
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR

; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-109

Query Match 11.0%; Score 13; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQRTQQQ 97
Db 11 KVLGLLQRTQQQ 23

RESULT 36

US-09-112-206-109
; Sequence 109, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:

; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-109

Query Match 11.0%; Score 13; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQRTQQQ 97
Db 11 KVLGLLQRTQQQ 23

RESULT 37

US-09-576-824A-106
; Sequence 106, Application US/09576824A

; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: modified site
; NAME/KEY: VARIANT
; LOCATION: (24)
; OTHER INFORMATION: modified site
; US-09-576-824A-106

Query Match 11.0%; Score 13; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQRTQQQ 97
Db 11 KVLGLLQRTQQQ 23

RESULT 38
US-09-680-497-109
; Sequence 109, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 109:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: peptide
US-09-680-497-109

Query Match 11.0%; Score 13; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLLQRTQQQ 97
Db 11 KVLGLLQRTQQQ 23

RESULT 39
US-09-288-391-26
; Sequence 26, Application US/09288391
; Patent No. 6251583
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Malcolm, Bruce
; APPLICANT: Beyer, Brian
; APPLICANT: Njoroge, George
; APPLICANT: Durkin, James
; APPLICANT: Windsor, William
; TITLE OF INVENTION: No. 6251583el Peptide Substrates for HCV NS3 Protease Assay
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/288,391
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: IN0829P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-288-391-26

Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 7 GGVLAALAAAYCLS 19

RESULT 40
US-10-104-966-7
; Sequence 7, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren

```
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: METHODS OF USE THEREOF
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS4A protein sequence
US-10-104-966-7
```

```
Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 7 GGVLAALAAAYCLS 19
```

```
RESULT 41
US-09-929-955-7
; Sequence 7, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS4A protein sequence
US-09-929-955-7
```

```
Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 7 GGVLAALAAAYCLS 19
```

```
RESULT 42
US-09-929-955-30
; Sequence 30, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
```

```
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS4A peptide
US-09-929-955-30
```

```
Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 7 GGVLAALAAAYCLS 19
```

```
RESULT 43
US-09-930-591-13
; Sequence 13, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS4A peptide
US-09-930-591-13
```

```
Query Match 11.0%; Score 13; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 7 GGVLAALAAAYCLS 19
```

```
RESULT 44
US-08-444-818-12
; Sequence 12, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
```

NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 135 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-12

Query Match 11.0%; Score 13; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 6.9e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 96 GGVLAALAAAYCLS 108

RESULT 45
US-08-336-553A-1
Sequence 1, Application US/08336553A
Patent No. 6054264
GENERAL INFORMATION:
APPLICANT: CHIEN, DAVID Y.
ADDRESSEE: KUO, GEORGE
TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993

ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-336-553A-1

Query Match 11.0%; Score 13; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 72 GGVLAALAAAYCLS 84

RESULT 46
US-08-439-157-1
Sequence 1, Application US/08439157
Patent No. 6416944
GENERAL INFORMATION:
APPLICANT: KUO, DAVID Y.
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,157
FILING DATE: 11-MAY-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553A
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-439-157-1

Query Match 11.0%; Score 13; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 72 GGVLAALAAAYCLS 84

RESULT 47

US-09-437-895-1
; Sequence 1, Application US/09437895
; Patent No. 6416946
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; Kuo, George
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/437,895
; FILING DATE: 09-No. 6416946-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 150 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-437-895-1

Query Match 11.0%; Score 13; DB 2; Length 150;
Best Local Similarity 100.0%; Pred. No. 7.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 72 GGVLAALAAAYCLS 84

RESULT 48

US-08-444-818-14
; Sequence 14, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.

; TITLE OF INVENTION: NANEV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-444-818-14

Query Match 11.0%; Score 13; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 96 GGVLAALAAAYCLS 108

RESULT 49

US-08-483-695-7
; Sequence 7, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:


```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-483-695-7

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 50
US-08-483-695-45
; Sequence 45, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-965-285-7

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 51
US-07-965-285-7
; Sequence 7, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-965-285-7

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 52
US-07-965-285-45
; Sequence 45, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-483-695-45

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 51
US-07-965-285-7
; Sequence 7, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-965-285-7

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 52
US-07-965-285-45
; Sequence 45, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-965-285-7

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 52
US-07-965-285-45
; Sequence 45, Application US/08483695
; Patent No. 5866139
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremendorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,695
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/965,285
; FILING DATE: 18-MAR-1993
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
```

```
; Sequence 45, Application US/07965285
; Patent No. 5879904
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-965-285-45

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 53
US-08-487-231-7
; Sequence 7, Application US/08487231
; Patent No. 5919454
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
```

```
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,231
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,285
; FILING DATE: 18-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-487-231-7

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 210 GGVLAALAAAYCLS 222

RESULT 54
US-08-487-231-45
; Sequence 45, Application US/08487231
; Patent No. 5919454
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Krensdorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; TITLE OF INVENTION: Applications
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Parabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,231
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/965,285
; FILING DATE: 18-MAR-1993
```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91 06 882
; FILING DATE: 06-JUN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-02000
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-487-231-45

Query Match 11.0%; Score 13; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLIS 30
Db 210 GGVLAALAAYCLIS 222

RESULT 55
US-09-201-912-7
; Sequence 7, Application US/09201912
; Patent No. 6210962
; GENERAL INFORMATION:
; APPLICANT: Brechot, Christian
; APPLICANT: Kremadorf, Dina
; APPLICANT: Porchon, Colette
; TITLE OF INVENTION: Nucleotide and Peptide Sequences of a
; TITLE OF INVENTION: Hepatitis C Virus Isolate, Diagnostic and Therapeutic
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/201,912
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,285
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05286-0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 313 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-201-912-45

Query Match 11.0%; Score 13; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLIS 30
Db 210 GGVLAALAAYCLIS 222

RESULT 57
US-08-867-611-37
; Sequence 37, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G

```

APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAWSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/646,757
FILING DATE:
APPLICATION NUMBER: US/08/179,896
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-611-37

Query Match 11.0%; Score 13; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCLS 30
Db 320 GGVLAALAAAYCLS 332

RESULT 58

US-09-690-359-37
Sequence 37, Application US/09690359
Patent No. 6593083
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
CASEY, JAMES M
DAILEY, STEPHEN H
DAWSON, GEORGE J
GUTIERREZ, ROBIN A
LESNIEWSKI, RICHARD R
STEWART, JAMES L
RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/690,359
FILING DATE: 17-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/867,611
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: US/08/646,757
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/179,896
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 342 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 37:
US-09-690-359-37

Query Match 11.0%; Score 13; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCLS 30
Db 320 GGVLAALAAAYCLS 332

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 59
US-08-867-611-38
; Sequence 38, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL USA
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/748,565
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELEPHONE: 708-937-9556
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-867-611-38

Query Match 11.0%; Score 13; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.00016;

Qy 18 GGVLAALAAAYCLS 30
|||||
Db 322 GGVLAALAAAYCLS 334
|||||
RESULT 60
US-09-690-359-38
; Sequence 38, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <unknown>
; APPLICATION NUMBER: US/07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US/07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US/07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US/07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US/07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 38:
US-09-690-359-38

Query Match 11.0%; Score 13; DB 2; Length 344;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 322 GGVLAALAAAYCLS 334

RESULT 61
US-08-867-611-39
Sequence 39, Application US/08867611
Patent No. 6172189
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAMSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
TITLE OF INVENTION: ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,757
FILING DATE:
APPLICATION NUMBER: US/08/179,896
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-611-39
Query Match 11.0%; Score 13; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 330 GGVLAALAAAYCLS 342
RESULT 62
US-09-690-359-39
Sequence 39, Application US/09690359
Patent No. 6593083
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAMSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
TITLE OF INVENTION: ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/690,359
FILING DATE: 17-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/867,611
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: US/08/646,757
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/179,896
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 39:

SEQUENCE CHARACTERISTICS:
LENGTH: 352 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 39:
US-09-690-359-39

Query Match 11.0%; Score 13; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 330 GGVLAALAAYCLS 342
|||||

RESULT 63
US-08-867-611-40
; Sequence 40, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E

REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556

INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 357 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-611-40

Query Match 11.0%; Score 13; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 335 GGVLAALAAYCLS 347
|||||

RESULT 64
US-09-690-359-40
; Sequence 40, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991

```

; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 40:
US-09-690-359-40

Query Match 11.0%; Score 13; DB 2; Length 357;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 335 GGVLAALAAAYCLS 347

RESULT 65
US-08-850-328-4
; Sequence 4, Application US/08850328
; Patent No. 637986
; GENERAL INFORMATION:
; APPLICANT: TAKAHAMA, Y.
; APPLICANT: SHIRASHI, J.
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS
; TITLE OF INVENTION: C VIRUS INFECTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,328
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maye, Thomas D
; REGISTRATION NUMBER: 34,524
; REFERENCE/DOCKET NUMBER: 32273-20004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 360 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-850-328-4

Query Match 11.0%; Score 13; DB 2; Length 360;
```

```

Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 24 GGVLAALAAAYCLS 36

RESULT 66
US-08-867-611-41
; Sequence 41, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAMSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```



```
;
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-867-611-41

Query Match 11.0%; Score 13; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 340 GGVLAALAAYCLS 352

RESULT 67
US-09-690-359-41
; Sequence 41, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; DESAI, SURESH M
; CASEY, JAMES M
; DAILEY, STEPHEN H
; DAWSON, GEORGE J
; GUTIERREZ, ROBIN A
; LESNIEWSKI, RICHARD R
; STEWART, JAMES L
; RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 362 amino acids
```

```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-690-359-41

Query Match 11.0%; Score 13; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 340 GGVLAALAAYCLS 352

RESULT 68
US-08-867-611-43
; Sequence 43, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; DESAI, SURESH M
; CASEY, JAMES M
; DAILEY, STEPHEN H
; DAWSON, GEORGE J
; GUTIERREZ, ROBIN A
; LESNIEWSKI, RICHARD R
; STEWART, JAMES L
; RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 363 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-867-611-43

Query Match 11.0%; Score 13; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 341 GGVLAALAAAYCLS 353

RESULT 69

US-09-690-359-43

; Sequence 43, Application US/09690359

; Patent No. 6593083

; GENERAL INFORMATION:

; APPLICANT: DEVARE, SUSHIL G
; DESAI, SURESH M
; CASEY, JAMES M
; DAILEY, STEPHEN H
; DAWSON, GEORGE J
; GUTIERREZ, ROBIN A
; LESNIEWSKI, RICHARD R
; STEWART, JAMES L
; SUPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT

ANTIGENS

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES

STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

CITY: ABBOTT PARK

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

APPLICANT APPLICATION DATA:

APPLICATION NUMBER: US/09/690,359

FILING DATE: 17-Oct-2000

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/867,611

FILING DATE: 02-JUN-1997

APPLICATION NUMBER: US/08/646,757

FILING DATE: <unknown>

APPLICATION NUMBER: US/08/179,896

FILING DATE: <unknown>

APPLICATION NUMBER: US 07/572,822

FILING DATE: 24-AUG-1990

APPLICATION NUMBER: US 07/614,069

FILING DATE: 07-NOV-1990

APPLICATION NUMBER: US 07/748,561

FILING DATE: 21-AUG-1991

APPLICATION NUMBER: US 07/748,565

FILING DATE: 21-AUG-1991

APPLICATION NUMBER: US 07/748,566

FILING DATE: 21-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: FORENBSKI, PRISCILLA E

; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834, US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 363 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-690-359-43

Query Match 11.0%; Score 13; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 341 GGVLAALAAAYCLS 353

RESULT 70

US-08-867-611-44

; Sequence 44, Application US/08867611

; Patent No. 6172189

; GENERAL INFORMATION:

; APPLICANT: DEVARE, SUSHIL G

; APPLICANT: DESAI, SURESH M

; APPLICANT: CASEY, JAMES M

; APPLICANT: DAILEY, STEPHEN H

; APPLICANT: DAWSON, GEORGE J

; APPLICANT: GUTIERREZ, ROBIN A

; APPLICANT: LESNIEWSKI, RICHARD R

; APPLICANT: STEWART, JAMES L

; APPLICANT: SUPRECHT, KEVIN R

; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT

ANTIGENS

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES

STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

CITY: ABBOTT PARK

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

APPLICANT APPLICATION DATA:

APPLICATION NUMBER: US/08/867,611

FILING DATE: 02-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,757

FILING DATE:

APPLICATION NUMBER: US/08/179,896

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/572,822

FILING DATE: 24-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/614,069

FILING DATE: 07-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/748,561

FILING DATE: 21-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/748,565

;; FILING DATE: 21-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,566
;; FILING DATE: 21-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834.US.P6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 44:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 364 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-867-611-44

Query Match 11.0%; Score 13; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 342 GGVLAALAAAYCLS 354

RESULT 71

US-09-690-359-44
; Sequence 44, Application US/09690359
; Patent No. 6593083

GENERAL INFORMATION:

;; APPLICANT: DEVARE, SUSHIL G
;; DESAI, SURESH M
;; CASEY, JAMES M
;; DAILEY, STEPHEN H
;; DAWSON, GEORGE J
;; GUTIERREZ, ROBIN A
;; LESNIEWSKI, RICHARD R
;; STEWART, JAMES L
;; RUPPRECHT, KEVIN R
;; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT ANTIGENS

;; NUMBER OF SEQUENCES: 59

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: ABBOTT LABORATORIES

;; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

;; CITY: ABBOTT PARK

;; STATE: IL

;; COUNTRY: USA

;; ZIP: 60064-3500

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent In Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/690,359

;; FILING DATE: 17-Oct-2000

;; CLASSIFICATION: <Unknown>

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 08/867,611

;; FILING DATE: 02-JUN-1997

;; APPLICATION NUMBER: US/08/646,757

;; FILING DATE: <Unknown>

;; APPLICATION NUMBER: US/08/179,896

;; FILING DATE: <Unknown>

;; APPLICATION NUMBER: US 07/572,822

;; FILING DATE: 24-AUG-1990

;; APPLICATION NUMBER: US 07/614,069

;; FILING DATE: 07-NOV-1990

;; APPLICATION NUMBER: US 07/748,561
;; FILING DATE: 21-AUG-1991
;; APPLICATION NUMBER: US 07/748,565
;; FILING DATE: 21-AUG-1991
;; APPLICATION NUMBER: US 07/748,566
;; FILING DATE: 21-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834.US.P6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 44:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 364 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-690-359-44

Query Match 11.0%; Score 13; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 342 GGVLAALAAAYCLS 354

RESULT 72

US-08-867-611-42

;; Sequence 42, Application US/08867611

;; Patent No. 6172189

;; GENERAL INFORMATION:

;; APPLICANT: DEVARE, SUSHIL G

;; APPLICANT: DESAI, SURESH M

;; APPLICANT: CASEY, JAMES M

;; APPLICANT: DAILEY, STEPHEN H

;; APPLICANT: DAWSON, GEORGE J

;; APPLICANT: GUTIERREZ, ROBIN A

;; APPLICANT: LESNIEWSKI, RICHARD R

;; APPLICANT: STEWART, JAMES L

;; APPLICANT: RUPPRECHT, KEVIN R

;; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT

;; TITLE OF INVENTION: ANTIGENS

;; NUMBER OF SEQUENCES: 59

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: ABBOTT LABORATORIES

;; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

;; CITY: ABBOTT PARK

;; STATE: IL

;; COUNTRY: USA

;; ZIP: 60064-3500

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent In Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/867,611

;; FILING DATE: 02-JUN-1997

;; CLASSIFICATION: 435

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US/08/646,757

;; FILING DATE:

;; APPLICATION NUMBER: US/08/179,896

;; FILING DATE:

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: US 07/572,822

;; FILING DATE: 24-AUG-1990

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-9365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-867-611-42

Query Match 11.0%; Score 13; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 343 GGVLAALAAAYCLS 355

RESULT 73
US-09-690-359-42
; Sequence 42, Application US/09690359
; Patent No. 6591083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; DESAI, SURESH M
; CASEY, JAMES M
; DAILEY, STEPHEN H
; DAWSON, GEORGE J
; GUTIERREZ, ROBIN A
; LESNIEWSKI, RICHARD R
; STEWART, JAMES L
; RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690.359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
```

```
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 365 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 42:
US-09-690-359-42

Query Match 11.0%; Score 13; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 0.00017;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 343 GGVLAALAAAYCLS 355

RESULT 74
US-08-444-818-68
; Sequence 68, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
```

TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 382 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-68

Query Match 11.0%; Score 13; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 0.00018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLIS 30
|||||
Db 110 GGVLAALAAYCLIS 122

RESULT 75

US-08-444-818-20
Sequence 20, Application US/08444818
Patent No. 6150087

GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 460 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-20

Query Match 11.0%; Score 13; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.00021;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLIS 30
|||||
Db 194 GGVLAALAAYCLIS 206

RESULT 76

US-08-867-611-47

Sequence 47, Application US/08867611
Patent No. 6172189
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAWSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,757
FILING DATE:
APPLICATION NUMBER: US/08/179,896
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMBSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 592 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-611-47

Query Match 11.0%; Score 13; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLIS 30
|||||

Db 320 GGVLAALAAAYCLS 332

RESULT 77

US-09-690-359-47
; Sequence 47, Application US/09690359
; Patent No. 6593083

GENERAL INFORMATION:

APPLICANT: DEVARE, SUSHIL G
DESAL, SURESH M
CASEY, JAMES M
DAILEY, STEPHEN H
DAWSON, GEORGE J
GUTIERREZ, ROBIN A
LESNIEWSKI, RICHARD R
STEWART, JAMES L
RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
ANTIGENS

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/690.359
FILING DATE: 17-Oct-2000
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/867,611
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: US/08/646,757
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/179,896
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:
LENGTH: 592 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-09-690-359-47

Query Match

Best Local Similarity 11.0%; Score 13; DB 2; Length 592;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 320 GGVLAALAAAYCLS 332

RESULT 78

US-08-867-611-48
; Sequence 48, Application US/08867611
; Patent No. 6172189

GENERAL INFORMATION:

APPLICANT: DEVARE, SUSHIL G
DESAL, SURESH M
CASEY, JAMES M
DAILEY, STEPHEN H
DAWSON, GEORGE J
GUTIERREZ, ROBIN A
LESNIEWSKI, RICHARD R
STEWART, JAMES L
RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
ANTIGENS
NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,757
FILING DATE: US/08/179,896
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-867-611-48

Query Match 11.0%; Score 13; DB 2; Length 594;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 322 GGVLAALAAAYCLS 334

RESULT 79

US-09-690-359-48
; Sequence 48, Application US/09690359
; Patent No. 6593083

GENERAL INFORMATION:

APPLICANT: DEVARE, SUSHIL G
DESAI, SURESH M
CASEY, JAMES M
DAILEY, STEPHEN H
DAWSON, GEORGE J
GUTIERREZ, ROBIN A
LESNIEWSKI, RICHARD R
STEWART, JAMES L
RUPPRECHT, KEVIN R

TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT ANTIGENS

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSER: ABBOTT LABORATORIES

STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

CITY: ABBOTT PARK

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/690,359

FILING DATE: 17-Oct-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/867,611

FILING DATE: 02-JUN-1997

APPLICATION NUMBER: US/08/646,757

FILING DATE: <Unknown>

APPLICATION NUMBER: US/08/179,896

FILING DATE: <Unknown>

APPLICATION NUMBER: US/07/572,822

FILING DATE: 24-AUG-1990

APPLICATION NUMBER: US/07/614,069

FILING DATE: 07-NOV-1990

APPLICATION NUMBER: US/07/748,561

FILING DATE: 21-AUG-1991

APPLICATION NUMBER: US/07/748,565

FILING DATE: 21-AUG-1991

APPLICATION NUMBER: US/07/748,566

FILING DATE: 21-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: FOREMSKI, PRISCILLA E

REGISTRATION NUMBER: 33,207

REFERENCE/DOCKET NUMBER: 4834.US.P6

TELEPHONE: 708-937-6365

TELEFAX: 708-937-9556

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 594 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 48:

US-09-690-359-48

Query Match 11.0%; Score 13; DB 2; Length 594;

Best Local Similarity 100.0%; Pred. No. 0.00027;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

Db 322 GGVLAALAAAYCLS 334

RESULT 80

US-08-867-611-16

; Sequence 16, Application US/08867611

; Patent No. 6172189

GENERAL INFORMATION:

APPLICANT: DEVARE, SUSHIL G
DESAI, SURESH M
CASEY, JAMES M
DAILEY, STEPHEN H
DAWSON, GEORGE J
GUTIERREZ, ROBIN A
LESNIEWSKI, RICHARD R
STEWART, JAMES L
RUPPRECHT, KEVIN R

TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT

NUMBER OF SEQUENCES: 59

CORRESPONDENCE ADDRESS:

ADDRESSER: ABBOTT LABORATORIES

STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2

CITY: ABBOTT PARK

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/867,611

FILING DATE: 02-JUN-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,757

FILING DATE:

APPLICATION NUMBER: US/08/179,896

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/572,822

FILING DATE: 24-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/614,069

FILING DATE: 07-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/748,561

FILING DATE: 21-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/748,565

FILING DATE: 21-AUG-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/748,566

FILING DATE: 21-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: FOREMSKI, PRISCILLA E

REGISTRATION NUMBER: 33,207

REFERENCE/DOCKET NUMBER: 4834.US.P6

TELEPHONE: 708-937-6365

```
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 597 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-867-611-16

Query Match      11.0%; Score 13; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      320 GGVLAALAAAYCLS 332

RESULT 81
US-09-690-359-16
; Sequence 16, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; CASEY, JAMES M
; DAILEY, STEPHEN H
; DAWSON, GEORGE J
; GUTIERREZ, ROBIN A
; LESNIEWSKI, RICHARD R
; STEWART, JAMES L
; RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 597 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-690-359-16

Query Match      11.0%; Score 13; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      320 GGVLAALAAAYCLS 332

RESULT 82
PCT-US92-06965A-21
; Sequence 21, Application PC/TUS9206965A
; GENERAL INFORMATION:
; APPLICANT: DEVARE, S.
; APPLICANT: DESAI, S.
; TITLE OF INVENTION: HCV SYNTHETIC PEPTIDE FROM NS1 REGION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: ILLINOIS
; COUNTRY: U.S.
; ZIP: 60065-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06965A
; FILING DATE: 19920821
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834PC.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 597 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; MOLECULE TYPE: protein
PCT-US92-06965A-21

Query Match      11.0%; Score 13; DB 4; Length 597;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      320 GGVLAALAAAYCLS 332

RESULT 83
US-08-867-611-18
; Sequence 18, Application US/08867611
; Patent No. 6172189
```


GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAWSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESS: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,757
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US/08/179,896
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US/07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US/07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-9556
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 599 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-867-611-18

Query Match 11.0%; Score 13; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
|||||
Db 322 GGVLAALAAYCLS 334

RESULT 84
US-09-690-359-18
Sequence 18, Application US/09690359
Patent No. 6593083
GENERAL INFORMATION:
APPLICANT: DEVARE, SUSHIL G
APPLICANT: DESAI, SURESH M
APPLICANT: CASEY, JAMES M
APPLICANT: DAILEY, STEPHEN H
APPLICANT: DAWSON, GEORGE J
APPLICANT: GUTIERREZ, ROBIN A
APPLICANT: LESNIEWSKI, RICHARD R
APPLICANT: STEWART, JAMES L
APPLICANT: RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESS: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/690,359
FILING DATE: 17-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/867,611
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: US/08/646,757
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/179,896
FILING DATE: <Unknown>
APPLICATION NUMBER: US/07/572,822
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US/07/614,069
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US/07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US/07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US/07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 599 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-690-359-18

Query Match 11.0%; Score 13; DB 2; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
|||||
Db 322 GGVLAALAAYCLS 334

RESULT 85
PCT-US92-06965A-23
; Sequence 23, Application PC/TUS9206965A
; GENERAL INFORMATION:
; APPLICANT: DEVARE, S.
; APPLICANT: DESAI, S.
; APPLICANT: DAILEY, S.
; TITLE OF INVENTION: HCV SYNTHETIC PEPTIDE FROM NS1 REGION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: ILLINOIS
; COUNTRY: U.S.
; ZIP: 60065-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/06965A
; FILING DATE: 19920821
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834PC.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 599 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-06965A-23

Query Match 11.0%; Score 13; DB 4; Length 599;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 322 GGVLAALAAAYCLS 334

RESULT 86
US-08-867-611-49
; Sequence 49, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL

COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,611
FILING DATE: 02-JUN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US/08/646,757
FILING DATE:
APPLICATION NUMBER: US/08/179,896
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 613 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-867-611-49

Query Match 11.0%; Score 13; DB 2; Length 613;
Best Local Similarity 100.0%; Pred. No. 0.00028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 341 GGVLAALAAAYCLS 353

RESULT 87
US-09-690-359-49
; Sequence 49, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; NUMBER OF SEQUENCES: 59
; ANTIGENS

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US/07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US/07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US/07/748,565
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US/07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 613 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 49:
US-09-690-359-49

Query Match 11.0%; Score 13; DB 2; Length 613;
Best Local Similarity 100.0%; Pred.No. 0.00028; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCLS 30
Db 341 GGVLAALAAYCLS 353

RESULT 88
US-09-881-239-3
; Sequence 3, Application US/09881239
; Patent No. 6630298
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
; FILE REFERENCE: 2302-16073 / PP16073.003
; CURRENT APPLICATION NUMBER: US/09/881,239

; CURRENT FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
US-09-881-239-3

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred.No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCLS 30
Db 639 GGVLAALAAYCLS 651

RESULT 89
US-09-881-654-2
; Sequence 2, Application US/09881654
; Patent No. 6632601
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PP17039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
US-09-881-654-2

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred.No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCLS 30
Db 639 GGVLAALAAYCLS 651

RESULT 90
US-10-637-323-2
; Sequence 2, Application US/10637323
; Patent No. 6797809
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKE, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica

; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / P17039.002
; CURRENT APPLICATION NUMBER: US/10/637,323
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/09/881,654
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/280,867
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
US-10-637-323-2

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 639 GGVLAALAAAYCLS 651

RESULT 91
US-09-929-955-17
; Sequence 17, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS3/4A peptide
US-09-929-955-17

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 639 GGVLAALAAAYCLS 651

RESULT 92
US-09-929-955-31
; Sequence 31, Application US/09929955
; Patent No. 6858590

; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-31

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 639 GGVLAALAAAYCLS 651

RESULT 93
US-09-929-955-32
; Sequence 32, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-32

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
| | | | | | | | | |
Db 639 GGVLAALAAAYCLS 651

RESULT 94
US-09-929-955-43
; Sequence 43, Application US/09929955

; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-43

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
| | | | | | | | | | | | | | | | | | | | | |
Db 639 GGVLAAALAAAYCLS 651

RESULT 95
US-09-929-955-44
; Sequence 44, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-44

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
| | | | | | | | | | | | | | | | | | | | | |
Db 639 GGVLAAALAAAYCLS 651

RESULT 96
US-09-929-955-45

; Sequence 45, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-45

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
| | | | | | | | | | | | | | | | | | | | | |
Db 639 GGVLAAALAAAYCLS 651

RESULT 97
US-09-929-955-46
; Sequence 46, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-46

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
| | | | | | | | | | | | | | | | | | | | | |
Db 639 GGVLAAALAAAYCLS 651

RESULT 98

```
US-09-929-955-47
; Sequence 47, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-47

Query Match      11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      639 GGVLAALAAAYCLS 651
|||||

RESULT 99
US-09-929-955-48
; Sequence 48, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-48

Query Match      11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      639 GGVLAALAAAYCLS 651
|||||

US-09-929-955-49
; Sequence 49, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-49

Query Match      11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      639 GGVLAALAAAYCLS 651
|||||

RESULT 100
US-09-929-955-49
; Sequence 49, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 49
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-929-955-49

Query Match      11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      639 GGVLAALAAAYCLS 651
|||||

RESULT 101
US-09-930-591-2
; Sequence 2, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 03/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS3/4A peptide
US-09-930-591-2

Query Match      11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCLS 30
Db      639 GGVLAALAAAYCLS 651
|||||
```

RESULT 102

US-09-930-591-3
; Sequence 3, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
|||||
Db 639 GGVLAALAAAYCLS 651

RESULT 103

US-09-930-591-4
; Sequence 4, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
|||||
Db 639 GGVLAALAAAYCLS 651

RESULT 104

US-09-930-591-4

US-09-930-591-5
; Sequence 5, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
|||||
Db 639 GGVLAALAAAYCLS 651

RESULT 105

US-09-930-591-6
; Sequence 6, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
|||||
Db 639 GGVLAALAAAYCLS 651

RESULT 106

US-09-930-591-7

; Sequence 7, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP 028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-930-591-7

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 107
US-09-930-591-8
; Sequence 8, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP 028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-930-591-8

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 108
US-09-930-591-9
; Sequence 9, Application US/09930591

; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP 028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-930-591-9

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 109
US-09-930-591-10
; Sequence 10, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP 028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-930-591-10

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 639 GGVLAALAAAYCLS 651

RESULT 110
US-09-930-591-11
; Sequence 11, Application US/09930591
; Patent No. 6960569

GENERAL INFORMATION:
APPLICANT: Matti Sallberg
TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
FILE REFERENCE: NS3/4A FUSION GENE
CURRENT APPLICATION NUMBER: US/09/930,591
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 686
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A
US-09-930-591-11

Query Match 11.0%; Score 13; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 0.00031; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLIS 30
|||||
Db 639 GGVLAALAAAYCLIS 651

RESULT 111
US-09-881-239-1
Sequence 1, Application US/09881239
Patent No. 6630298
GENERAL INFORMATION:
APPLICANT: CHIEN, David Y.
APPLICANT: ARCANGEL, Phillip
APPLICANT: TANDESKE, Laura
APPLICANT: GEORGE-NASCIEMENTO, Carlos
APPLICANT: COIT, Doris
APPLICANT: MEDINA-SELBY, Angelica
TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
FILE REFERENCE: 2302-16073 / PP16073.003
CURRENT APPLICATION NUMBER: US/09/881,239
CURRENT FILING DATE: 2001-06-14
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 728
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:
OTHER INFORMATION: representative NS3/4a conformational antigen
US-09-881-239-1

Query Match 11.0%; Score 13; DB 2; Length 728;
Best Local Similarity 100.0%; Pred. No. 0.00033; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLIS 30
|||||
Db 681 GGVLAALAAAYCLIS 693

RESULT 112
US-08-444-818-148
Sequence 148, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.

APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 739 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-148

Query Match 11.0%; Score 13; DB 2; Length 739;
Best Local Similarity 100.0%; Pred. No. 0.00033; Mismatches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLIS 30
|||||
Db 473 GGVLAALAAAYCLIS 485

RESULT 113
US-08-444-818-30
Sequence 30, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590

```
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-444-818-30

Query Match 11.0%; Score 13; DB 2; Length 859;
Best Local Similarity 100.0%; Pred. No. 0.00038;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 399 GGVLAALAAAYCLS 411

RESULT 114
US-08-667-611-52
; Sequence 52, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: POREBSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 971 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-867-611-52

Query Match 11.0%; Score 13; DB 2; Length 971;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 699 GGVLAALAAAYCLS 711

RESULT 115
US-09-690-359-52
; Sequence 52, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
```

;; FILING DATE: 07-NOV-1990
;; APPLICATION NUMBER: US 07/748,561
;; FILING DATE: 21-AUG-1991
;; APPLICATION NUMBER: US 07/748,565
;; FILING DATE: 21-AUG-1991
;; APPLICATION NUMBER: US 07/748,566
;; FILING DATE: 21-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834.US.P6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 52:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 971 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-690-359-52

Query Match 11.0%; Score 13; DB 2; Length 971;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 699 GGVLAALAAYCLS 711

RESULT 116
US-08-867-611-53
; Sequence 53, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822

;; FILING DATE: 24-AUG-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/614,069
;; FILING DATE: 07-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,561
;; FILING DATE: 21-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,565
;; FILING DATE: 21-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,566
;; FILING DATE: 21-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834.US.P6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 53:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 973 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-867-611-53

Query Match 11.0%; Score 13; DB 2; Length 973;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 701 GGVLAALAAYCLS 713

RESULT 117
US-09-690-359-53
; Sequence 53, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611

; FILING DATE: 02-JUN-1997
; APPLICATION NUMBER: US/08/646,757
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/179,896
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMBSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 973 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-09-690-359-53

Query Match 11.0%; Score 13; DB 2; Length 973;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
Db 701 GGVLAAALAAAYCLS 713

RESULT 118
US-08-867-611-54
; Sequence 54, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMBSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 992 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-867-611-54

Query Match 11.0%; Score 13; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCLS 30
Db 720 GGVLAAALAAAYCLS 732

RESULT 119
US-09-690-359-54
; Sequence 54, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/690,359
FILING DATE: 17-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/867,611
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: US/08/646,757
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/179,896
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-09-690-359-54

Query Match 11.0%; Score 13; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
DB 720 GGVLAALAAAYCLS 732

RESULT 120
US-07-910-760-12
Sequence 12, Application US/07910760
Patent No. 568384
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/910,760
FILING DATE: 07-JUL-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-910-760-12

Query Match 11.0%; Score 13; DB 1; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
DB 632 GGVLAALAAAYCLS 644

RESULT 121
US-08-440-519-12
Sequence 12, Application US/08440519
Patent No. 5712087
GENERAL INFORMATION:
APPLICANT: Houghton, Michael
APPLICANT: Choo, Qui-Lim
APPLICANT: Kuo, George
TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: P.O. Box 8097 (Int. Prop. R-440)
CITY: Emeryville
STATE: CA
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,519
FILING DATE: 12-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/910,760
FILING DATE: 07-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: 0101.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1021 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-440-519-12

Query Match 11.0%; Score 13; DB 1; Length 1021;

```
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 632 GGVLAALAAAYCLS 644

RESULT 122
US-08-440-549-12
; Sequence 12, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1021 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-549-12

Query Match 11.0%; Score 13; DB 2; Length 1021;
Best Local Similarity 100.0%; Pred. No. 0.00045;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 632 GGVLAALAAAYCLS 644

RESULT 123
US-08-444-818-66
; Sequence 66, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
```

```
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508) 359-3876
TELEFAX: (508) 359-3885
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 2261 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-66

Query Match 11.0%; Score 13; DB 2; Length 2261;
Best Local Similarity 100.0%; Pred. No. 0.00093;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1039 GGVLAALAAAYCLS 1051

RESULT 124
US-08-444-818-75
; Sequence 75, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 75:
SEQUENCE CHARACTERISTICS:
LENGTH: 2436 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-75

Query Match 11.0%; Score 13; DB 2; Length 2436;
Best Local Similarity 100.0%; Pred. No. 0.00099;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 1214 GGVLAALAAYCLS 1226

RESULT 125

US-08-444-818-89
Sequence 89, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (508)359-3876

TELEFAX: (508)359-3885

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:

LENGTH: 2772 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-444-818-89

Query Match 11.0%; Score 13; DB 2; Length 2772;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 1550 GGVLAALAAYCLS 1562

RESULT 126

US-08-466-975A-23
Sequence 23, Application US/08466975A
Patent No. 5910404
GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT J

APPLICANT: POLLET, DIRK

APPLICANT: MAERTENS, GEERT

APPLICANT: VAN HEUVERSWUN, HUGO

TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHUYE P.C.

STREET: 1100 NORTH GLEBE ROAD

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,975A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/391,671

FILING DATE:

APPLICATION NUMBER: US/07/920,286

FILING DATE: 14-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/EP91/02409

FILING DATE: 13-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 90124241.2

FILING DATE: 14-DEC-1990

ATTORNEY/AGENT INFORMATION:

NAME: SADOFF, B.J.

REGISTRATION NUMBER: 36,663

REFERENCE/DOCKET NUMBER: 1487-5

TELECOMMUNICATION INFORMATION:

TELEPHONE: 7038164000

TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 2894 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-466-975A-23

Query Match 11.0%; Score 13; DB 1; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 1664 GGVLAALAAYCLS 1676

RESULT 127

US-08-391-671A-23
Sequence 23, Application US/08391671A
Patent No. 5922532
GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT J

APPLICANT: POLLET, DIRK

APPLICANT: MAERTENS, GEERT

APPLICANT: VAN HEUVERSWUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671A
FILING DATE: 21-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 2894 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-391-671A-23

Query Match 11.0%; Score 13; DB 1; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 128
US-08-467-902A-23
Sequence 23, Application US/08467902A
Patent No. 6007982
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
APPLICANT: VAN HEUVERSWUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA

COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,902A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671
FILING DATE:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 2894 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-467-902A-23

Query Match 11.0%; Score 13; DB 2; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 129
US-09-275-265-23
Sequence 23, Application US/09275265
Patent No. 6287761
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
APPLICANT: VAN HEUVERSWUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/275,265
;; FILING DATE: 13-DEC-1991
;; CLASSIFICATION: EP 90124241.2
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/391,671
;; FILING DATE: 21-FEB-1995
;; APPLICATION NUMBER: US 07/920,286
;; FILING DATE: 14-OCT-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/EP91/02409
;; FILING DATE: 13-DEC-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: EP 90124241.2
;; FILING DATE: 14-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 7038164000
;; TELEFAX: 7038164100
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2894 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
US-09-275-265-23

Query Match 11.0%; Score 13; DB 2; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 1664 GGVLAALAAYCLS 1676

RESULT 130
US-09-941-611-23
; Sequence 23, Application US/09941611
; Patent No. 6576417
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,611
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: 1995-02-21

;; APPLICATION NUMBER: WO PCT/EP91/02409
;; FILING DATE: 13-DEC-1991
;; APPLICATION NUMBER: EP 90124241.2
;; FILING DATE: 14-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-5
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 7038164000
;; TELEFAX: 7038164100
;; INFORMATION FOR SEQ ID NO: 23:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2894 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-941-611-23

Query Match 11.0%; Score 13; DB 2; Length 2894;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
Db 1664 GGVLAALAAYCLS 1676

RESULT 131
US-10-044-995-23
; Sequence 23, Application US/10044995
; Patent No. 6872520
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/044,995
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5

```
;
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 7038164000
;   TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 23:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2894 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: peptide
;     HYPOTHETICAL: NO
;     ANTI-SENSE: NO
;     SEQUENCE DESCRIPTION: SEQ ID NO: 23:
;
; US-10-044-995-23
;
; Query Match 11.0%; Score 13; DB 2; Length 2894;
; Best Local Similarity 100.0%; Pred. No. 0.0012;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCLS 30
; Db 1664 GGVLAALAAAYCLS 1676
;
; RESULT 132
; US-08-443-260-3
; Sequence 3, Application US/08443260
; Patent No. 5942234
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDIUM, KENT B.
; APPLICANT: GERVAISE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,260
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154.006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2349
; OTHER INFORMATION: /note= "There is a heterogeneity at
;
; ;
; LOCATION: 11
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 176
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 334
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 603
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 848
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1117
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1276
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1454
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1471
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1877
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1948
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1949
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2021
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2349
; OTHER INFORMATION: /note= "There is a heterogeneity at
; FEATURE:
; NAME/KEY: Modified-site
```

FEATURE: NAME/KEY: Modified-site
LOCATION: 2385
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Phe or Tyr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 2386
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ala or Ser"
FEATURE: NAME/KEY: Modified-site
LOCATION: 2502
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Phe or Leu"
FEATURE: NAME/KEY: Modified-site
LOCATION: 2690
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Gly or Arg"
FEATURE: NAME/KEY: Modified-site
LOCATION: 2921
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Arg or Gly"
US-08-443-260-3

Query Match 11.0%; Score 13; DB 1; Length 2955;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

Db 1664 GGVLAALAAAYCLS 1676

RESULT 133

US-08-442-805A-3
Sequence 3, Application US/08442805A
Patent No. 6074846
GENERAL INFORMATION:
APPLICANT: RALSTON, ROBERT O.
APPLICANT: MARCUS, FRANK
APPLICANT: THUDUM, KENT B.
APPLICANT: GERVAISE, BARBARA A.
APPLICANT: HALL, JOHN A.
TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,805A
FILING DATE: 17-MAY-1995
CLASSIFICATION: 520
ATTORNEY/AGENT INFORMATION:
NAME: HARBIN, ALISA A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0154.005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:

LENGTH: 2955 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE: NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Arg or Lys"
FEATURE: NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Asn or Thr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 176
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ile or Thr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 334
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Met or Val"
FEATURE: NAME/KEY: Modified-site
LOCATION: 603
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ile or Leu"
FEATURE: NAME/KEY: Modified-site
LOCATION: 848
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Asn or Tyr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1114
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Pro or Ser"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1117
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1276
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Leu or Pro"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Cys or Tyr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1471
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1877
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Glu or Gly"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1948
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = His or Leu"
FEATURE: NAME/KEY: Modified-site
LOCATION: 1949

```

; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2021
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2349
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2385
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2396
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ala or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2502
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2690
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2921
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Gly"
;
; US-08-442-805A-3
;
; Query Match 11.0%; Score 13; DB 2; Length 2955;
; Best Local Similarity 100.0%; Pred. No. 0.0012;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCLS 30
; 1664 GGVLAALAAAYCLS 1676
;
; Db
;
; RESULT 134
; US-08-443-900A-3
; Sequence 3, Application US/08443900A
; Patent No. 6074852
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDUM, KENT B.
; APPLICANT: GERVASE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: US/08/443,900A
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Lys"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Asn or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 176
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 334
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Met or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 603
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ile or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 848
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Asn or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1114
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Pro or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1117
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1276
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Leu or Pro"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1454
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1471
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:

```

```

; NAME/KEY: Modified-site
; LOCATION: 1877
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Glu or Gly"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1948
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = His or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1949
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Cys or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2021
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2349
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2385
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2386
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ala or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2502
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Phe or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2690
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Gly or Arg"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2921
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Gly"
;
US-08-443-900A-3
Query Match 11.0%; Score 13; DB 2; Length 2955;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||
RESULT 135
US-08-444-818-124
; Sequence 124, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
;
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 124:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 9
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Lys or Arg"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 11
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Asn or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 176
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Ile or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 334
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Met or Val"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 603
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Leu or Ile"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 848
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Tyr or Asn"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1276
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Pro or Leu"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1454
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Cys or Tyr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1471
; OTHER INFORMATION: /note= "A heterogeneity exists at
```

```
;
; OTHER INFORMATION: Xaa which is either Thr or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1877
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Glu or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1948
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Leu or His"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1949
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Ser or Cys"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2021
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Gly or Val"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2349
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Thr or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2385
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Tyr or Phe"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2386
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Ser or Ala"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2502
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Leu or Phe"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2690
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Arg or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2921
; OTHER INFORMATION: /note= "A heterogeneity exists at
; OTHER INFORMATION: Xaa which is either Arg or Gly"
; US-08-444-818-124
;
; Query Match 11.0%; Score 13; DB 2; Length 2955;
; Best Local Similarity 100.0%; Pred. No. 0.0012;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCLS 30
; Db 1664 GGVLAALAAAYCLS 1676
;
; RESULT 136
; US-08-249-843-3
; Sequence 3, Application US/08249843
; Patent No. 6274148
; GENERAL INFORMATION:
; APPLICANT: RALSTON, ROBERT O.
; APPLICANT: MARCUS, FRANK
; APPLICANT: THUDIUM, KENT B.
; APPLICANT: GERVAISE, BARBARA A.
; APPLICANT: HALL, JOHN A.
; TITLE OF INVENTION: HEPATITIS C VIRUS ASIALOGLYCOPROTEINS
;
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/249,843
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HARBIN, ALISA A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0154.003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2955 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Arg or Lys"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Asn or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 176
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 334
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Met or Val"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 603
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ile or Leu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 848
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Asn or Tyr"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1114
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Pro or Ser"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1117
; OTHER INFORMATION: /note= "There is a heterogeneity at
; OTHER INFORMATION: this location; Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Modified-site
```

LOCATION: 1276
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Leu or Pro"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1454
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Cys or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1471
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1877
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Glu or Gly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1948
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = His or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1949
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Cys or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2021
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Gly or Val"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2349
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2385
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Phe or Tyr"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2386
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Ala or Ser"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2502
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Phe or Leu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2690
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Gly or Arg"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2921
OTHER INFORMATION: /note= "There is a heterogeneity at
OTHER INFORMATION: this location; Xaa = Arg or Gly"
US-08-249-843-3

Query Match 11.0%; Score 13; DB 2; Length 2955;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAAALAAAYCLS 30
Db 1664 GGVLAAALAAAYCLS 1676

RESULT 137
US-08-444-818-138
; Sequence 138, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: 777
; APPLICATION NUMBER: US/08/444.818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403.590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2995 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-444-818-138

Query Match 11.0%; Score 13; DB 2; Length 2995;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAAALAAAYCLS 30
Db 1664 GGVLAAALAAAYCLS 1676

RESULT 138
US-08-188-281B-1
; Sequence 1, Application US/08188281B
; Patent No. 5610009
; GENERAL INFORMATION:
; APPLICANT: WATANABE, SHINICHI
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: DESAI, SURESH M.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,281B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5521.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-188-281B-1

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 139
US-08-453-552-1
; Sequence 1, Application US/08453552
; Patent No. 5667992
; GENERAL INFORMATION:
; APPLICANT: CASEY, JAMES M.
; APPLICANT: BODE, SUZANNE L.
; APPLICANT: ZECK, BILLY J.
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: FRAIL, DONALD E.
; APPLICANT: DESAI, SURESH M.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5131.US.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-552-1

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 140
US-08-453-552-2
; Sequence 2, Application US/08453552
; Patent No. 5667992
; GENERAL INFORMATION:
; APPLICANT: CASEY, JAMES M.
; APPLICANT: BODE, SUZANNE L.
; APPLICANT: ZECK, BILLY J.
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: FRAIL, DONALD E.
; APPLICANT: DESAI, SURESH M.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5131.US.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-552-2

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 141
```

```
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-552-1

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 140
US-08-453-552-2
; Sequence 2, Application US/08453552
; Patent No. 5667992
; GENERAL INFORMATION:
; APPLICANT: CASEY, JAMES M.
; APPLICANT: BODE, SUZANNE L.
; APPLICANT: ZECK, BILLY J.
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: FRAIL, DONALD E.
; APPLICANT: DESAI, SURESH M.
; APPLICANT: DEVARE, SUSHIL G.
; TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5131.US.D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-453-552-2

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 141
```


US-08-440-103-36
; Sequence 36, Application US/08440103
; Patent No. 5670152
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440.103
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231.368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-103-36

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
|||
Db 1664 GGVLAALAAYCLS 1676

RESULT 142
US-08-440-542-36
; Sequence 36, Application US/08440542
; Patent No. 5670153
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440.542
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231.368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-542-36

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCLS 30
|||
Db 1664 GGVLAALAAYCLS 1676

RESULT 143
US-07-910-760-10
; Sequence 10, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; REFERENCE/DOCKET NUMBER: 0101.002
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/910.760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-910-760-10

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 144
US-08-440-519-10
; Sequence 10, Application US/08440519
; Patent No. 5712087
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; PRIOR APPLICATION NUMBER:
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 10:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-519-10

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 145
US-08-231-368-36
; Sequence 36, Application US/08231368
; Patent No. 5756312
; GENERAL INFORMATION:
```

```
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-231-368-36

Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 146
US-08-440-210-36
; Sequence 36, Application US/08440210
; Patent No. 5766845
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,210
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
```

TELEFAX: 708-937-9556

Matches 13; Conservative

;; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Phe"
;; FEATURE:
;; NAME/KEY: Duplication
;; LOCATION: 2690
;; OTHER INFORMATION: /note= "There exists a
;; OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
;; FEATURE:
;; NAME/KEY: Duplication
;; LOCATION: 2921
;; OTHER INFORMATION: /note= "There exists a
;; OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
;; FEATURE:
;; NAME/KEY: Duplication
;; LOCATION: 2996
;; OTHER INFORMATION: /note= "There exists a
;; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Pro"
US-08-833-678A-6
Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
|||||
Db 1664 GGVLAALAAAYCLS 1676
RESULT 150
US-08-811-566-20
; Sequence 20, Application US/08811566
; Patent No. 6127116
; GENERAL INFORMATION:
; APPLICANT: Rice, Charles et al.
; TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
; TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811.566
; FILING DATE: 03-MAR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1113-1-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: N-terminal
US-08-811-566-20
Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
|||||
Db 1664 GGVLAALAAAYCLS 1676
RESULT 151
US-08-444-818-177
; Sequence 177, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444.818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403.590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 177:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 9
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Lys or Arg"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 11
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Asn or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 176
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 334
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Met or Val"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 603
; OTHER INFORMATION: /note= "There exists a

```
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Ile"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 848
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Asn"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1114
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1117
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1276
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Leu"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1454
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Cys or Tyr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1471
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Thr or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1877
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Glu or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1948
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or His"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1949
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Cys"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2021
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Gly or Val"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2349
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Thr or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2385
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Phe"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2386
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Ala"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2502
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Phe"
; FEATURE:
; NAME/KEY: Duplication
```

```
; LOCATION: 2690
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2921
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2996
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Pro"
; US-08-444-818-177

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 152
US-09-014-416-1
; Sequence 1, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; US-09-014-416-1

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 153
US-09-014-416-5
; Sequence 5, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; 
```

```
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-014-416-5

Query Match      11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAAAYCLS 30
DB      1664 GGVLAALAAAYCLS 1676

RESULT 154
US-08-529-169A-6
; Sequence 6, Application US/08529169A
; Patent No. 6194140
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI-LIM
; APPLICANT: HAN, JANG
; APPLICANT: CHOE, JOONHO
; TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: Intellectual Property - R440, P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,169A
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 4325
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0100.005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3274
; TELEFAX: (510) 655-3542
; TELEX: n/a
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 9
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Lys or Arg"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 11
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Asn or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 176

; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ile or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 334
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Met or Val"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 603
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Ile"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 848
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Asn"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1114
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1117
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Thr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1276
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Leu"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1454
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Cys or Tyr"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1471
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Thr or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1877
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Glu or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1948
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or His"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 1949
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Cys"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2021
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Gly or Val"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2349
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Thr or Ser"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2385
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Phe"
; FEATURE:
```

```
; NAME/KEY: Duplication
; LOCATION: 2386
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Ala"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2502
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Phe"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2690
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2921
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
; FEATURE:
; NAME/KEY: Duplication
; LOCATION: 2996
; OTHER INFORMATION: /note= "There exists a
; OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Pro"
; US-08-529-169A-6
;
; Query Match 11.0%; Score 13; DB 2; Length 3011;
; Best Local Similarity 100.0%; Pred. No. 0.0012;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCLS 30
; Db 1664 GGVLAALAAAYCLS 1676
;
; RESULT 155
; US-09-388-874-2
; Sequence 2, Application US/09388874
; Patent No. 6284249
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/388,874
; CURRENT FILING DATE: 1999-09-02
; EARLIER APPLICATION NUMBER: PCT/FR98/00448
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 97/02,887
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Virus
; US-09-388-874-2
;
; Query Match 11.0%; Score 13; DB 2; Length 3011;
; Best Local Similarity 100.0%; Pred. No. 0.0012;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCLS 30
; Db 1664 GGVLAALAAAYCLS 1676
;
; RESULT 156
; US-09-046-604-36
; Sequence 36, Application US/09046604
; Patent No. 6303292
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
```

```
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,604
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-046-604-36
;
; Query Match 11.0%; Score 13; DB 2; Length 3011;
; Best Local Similarity 100.0%; Pred. No. 0.0012;
; Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCLS 30
; Db 1664 GGVLAALAAAYCLS 1676
;
; RESULT 157
; US-08-440-549-10
; Sequence 10, Application US/08440549
; Patent No. 6312889
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Lim
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```


APPLICANT: CHOE, JOONHO
TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: Intellectual Property - R440, P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/483,799
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/529,169
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0100.005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-3274
TELEFAX: (510) 655-3542
TELEX: n/a
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 3011 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Duplication
LOCATION: 9
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Lys or Arg"
FEATURE:
NAME/KEY: Duplication
LOCATION: 11
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Asn or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 176
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Ile or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 334
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Met or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 603
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Ile"
FEATURE:
NAME/KEY: Duplication
LOCATION: 848
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Asn"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1114
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Ser"

FEATURE:
NAME/KEY: Duplication
LOCATION: 1117
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Thr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1276
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Pro or Leu"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1454
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Cys or Tyr"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1471
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1877
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Glu or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1948
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or His"
FEATURE:
NAME/KEY: Duplication
LOCATION: 1949
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Cys"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2021
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Gly or Val"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2349
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Thr or Ser"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2385
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Tyr or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2386
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Ser or Ala"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2502
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Phe"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2690
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2921
OTHER INFORMATION: /note= "There exists a
OTHER INFORMATION: heterogeneity at this position - Xaa = Arg or Gly"
FEATURE:
NAME/KEY: Duplication
LOCATION: 2996

; OTHER INFORMATION: /note= "there exists a
; US-09-483-799-6 OTHER INFORMATION: heterogeneity at this position - Xaa = Leu or Pro"

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 161

US-09-916-359-2
; Sequence 2, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Virus
US-09-916-359-2

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 162

US-10-104-966-1
; Sequence 1, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus sequence
US-10-104-966-1

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 163

US-09-952-572-9
; Sequence 9, Application US/09952572
; Patent No. 6682909
; GENERAL INFORMATION:
; APPLICANT: HAWAII BIOTECHNOLOGY GROUP, Inc.
; APPLICANT: NAKANO, Bileen
; APPLICANT: CLEMENTS, David
; APPLICANT: HUMPHREYS, Tom
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITION OF HEPATITIS C
; FILE REFERENCE: HAWBIO1100
; CURRENT APPLICATION NUMBER: US/09/952,572
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: US 60/230,927
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-09-952-572-9

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

RESULT 164

US-09-929-955-1
; Sequence 1, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; PRIOR FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus sequence
US-09-929-955-1

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676
|||||

Db 1664 GGVLAALAAAYCLS 1676

RESULT 165

US-10-259-275-20

Sequence 20, Application US/10259275

Patent No. 6921634

GENERAL INFORMATION:

APPLICANT: Lemon, Stanley M.

APPLICANT: Yi, Minkyung

TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE

FILE REFERENCE: 265.0007 0120

CURRENT APPLICATION NUMBER: US/10/259,275

CURRENT FILING DATE: 2003-01-13

PRIOR APPLICATION NUMBER: US 60/171,909

PRIOR FILING DATE: 1999-12-23

PRIOR APPLICATION NUMBER: US 09/747,419

PRIOR FILING DATE: 2000-12-23

PRIOR APPLICATION NUMBER: US 60/325,236

PRIOR FILING DATE: 2001-09-27

PRIOR APPLICATION NUMBER: US 60/338,123

PRIOR FILING DATE: 2001-11-13

NUMBER OF SEQ ID NOS: 73

SOFTWARE: PatentIn version 3.0

SEQ ID NO 20

LENGTH: 3011

TYPE: PRT

ORGANISM: artificial

FEATURE:

OTHER INFORMATION: Polypeptide

US-10-259-275-20

Query Match 11.0%; Score 13; DB 2; Length 3011;

Best Local Similarity 100.0%; Pred. No. 0.0012;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

Db 1664 GGVLAALAAAYCLS 1676

RESULT 166

PCT-US91-02225-10

Sequence 10, Application PC/TUS9102225

GENERAL INFORMATION:

APPLICANT: HOUGHTON, MICHAEL

APPLICANT: CHOO, QUI-LIM

APPLICANT: KUO, GEORGE

TITLE OF INVENTION: COMBINATIONS OF HEPATITIS C VIRUS

TITLE OF INVENTION:

TITLE OF INVENTION: ANTIBODIES

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morrison & Foerster

STREET: 545 Middlefield Road, Suite 200

CITY: Menlo Park

STATE: CA

COUNTRY: USA

ZIP: 94025

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/02225

FILING DATE: 19910329

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CIOTTI, THOMAS E.

REGISTRATION NUMBER: 21,013

REFERENCE/DOCKET NUMBER: 2300-0101.44

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 327-7250

TELEFAX: (415) 327-2951

TELEX: 706141

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 3011 amino acids

TYPE: AMINO ACID

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Modified-site

LOCATION: 9

OTHER INFORMATION: /label= ArgorLys

OTHER INFORMATION: /note= "Amino acid 8 may be either an Arg or Lys"

FEATURE:

NAME/KEY: Modified-site

LOCATION: 11

OTHER INFORMATION: /note= "This amino acid may be either Asn or Thr."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 176

OTHER INFORMATION: /note= "This amino acid may be either Thr or Ile."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 334

OTHER INFORMATION: /note= "This amino acid may be either Val or Met."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 603

OTHER INFORMATION: /note= "This amino acid may be either Ile or Leu."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 848

OTHER INFORMATION: /note= "This amino acid may be either Asn or Tyr."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1114

OTHER INFORMATION: /note= "This amino acid may be either Pro or Ser."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1117

OTHER INFORMATION: /note= "This amino acid may be either Ser or Thr."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1276

OTHER INFORMATION: /note= "This amino acid may be either Leu or Pro."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1328

OTHER INFORMATION: /note= "This amino acid may be either Gly or Val."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1454

OTHER INFORMATION: /note= "This amino acid may be either Tyr or Cys."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1471

OTHER INFORMATION: /note= "This amino acid may be either Ser or Thr."

FEATURE:

NAME/KEY: Modified-site

LOCATION: 1877
OTHER INFORMATION: /note= "This amino acid may be
FEATURE: either Glu or Gly."
NAME/KEY: Modified-site
LOCATION: 1948
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either His or Leu."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1949
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Cys or Ser."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2021
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Gly or Val."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2349
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Thr or Ser."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2385
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Phe or Tyr."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2386
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Ser or Ala."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2502
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Leu or Phe."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2690
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Gly or Arg."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2921
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: either Gly or Arg."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2996
OTHER INFORMATION: /note= "This amino acid may be
OTHER INFORMATION: Leu or Pro."

Query Match 11.0%; Score 13; DB 4; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 1664 GGVLAALAAAYCLS 1676

RESULT 167
PCT-US93-00907-1
; Sequence 1, Application PC/TUS9300907
; GENERAL INFORMATION:
; APPLICANT: CASEY, JAMES M.
; APPLICANT: BODE, SUZANNE L.
; APPLICANT: ZECK, BILLY J.
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: FRAIL, DONALD E.

APPLICANT: DESAI, SURESH M.
APPLICANT: DEVARE, SUSHIL G.
TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00907
FILING DATE: 19930129
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: POREMSKI, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 5131.PC.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3011 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-00907-1

Query Match 11.0%; Score 13; DB 4; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 1664 GGVLAALAAAYCLS 1676

RESULT 168
PCT-US93-00907-2
; Sequence 2, Application PC/TUS9300907
; GENERAL INFORMATION:
; APPLICANT: CASEY, JAMES M.
; APPLICANT: BODE, SUZANNE L.
; APPLICANT: ZECK, BILLY J.
; APPLICANT: YAMAGUCHI, JULIE
; APPLICANT: FRAIL, DONALD E.
; APPLICANT: DESAI, SURESH M.
; APPLICANT: DEVARE, SUSHIL G.
TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

```
/ APPLICATION NUMBER: PCT/US93/00907
/ FILING DATE: 19930129
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: FOREMSKI, PRISCILLA E.
/ REGISTRATION NUMBER: 33,207
/ REFERENCE/DOCKET NUMBER: 5131.PC.01
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 708-937-6365
/ TELEFAX: 708-937-9556
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3011 amino acids
/ TYPE: AMINO ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US93-00907-2

Query Match 11.0%; Score 13; DB 4; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 169
PCT-US94-07280-1
/ Sequence 1, Application PC/TUS9407280
/ GENERAL INFORMATION:
/ APPLICANT: WATANABE, SHINICHI
/ APPLICANT: YAMAGUCHI, JULIE
/ APPLICANT: DESAI, SURESH M.
/ APPLICANT: DEVARE, SUSHIL G.
/ TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ STREET: ONE ABBOTT PARK ROAD
/ CITY: ABBOTT PARK
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60064-3500
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US94/07280
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: FOREMSKI, PRISCILLA E.
/ REGISTRATION NUMBER: 33,207
/ REFERENCE/DOCKET NUMBER: 5521.US.01
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 708-937-6365
/ TELEFAX: 708-938-2623
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3011 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US94-07280-1

Query Match 11.0%; Score 13; DB 4; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 171
US-08-811-566-2
/ Sequence 2, Application US/08811566
/ Patent No. 6127116
/ GENERAL INFORMATION:
/ APPLICANT: Rice, Charles et al.
/ TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
/ TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: David A. Jackson, Esq.
/ STREET: 411 Hackensack Ave, Continental Plaza, 4th
/ STREET: Floor
/ CITY: Hackensack
```

```
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 170
PCT-US95-01087-1
/ Sequence 1, Application PC/TUS9501087
/ GENERAL INFORMATION:
/ APPLICANT: WATANABE, SHINICHI
/ APPLICANT: YAMAGUCHI, JULIE
/ APPLICANT: DESAI, SURESH M.
/ APPLICANT: DEVARE, SUSHIL G.
/ TITLE OF INVENTION: MAMMALIAN EXPRESSION SYSTEMS FOR HCV
/ NUMBER OF SEQUENCES: 22
/ CORRESPONDENCE ADDRESS:
/ STREET: ONE ABBOTT PARK ROAD
/ CITY: ABBOTT PARK
/ STATE: IL
/ COUNTRY: USA
/ ZIP: 60064-3500
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/01087
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: FOREMSKI, PRISCILLA E.
/ REGISTRATION NUMBER: 33,207
/ REFERENCE/DOCKET NUMBER: 5521.US.01
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 708-937-6365
/ TELEFAX: 708-938-2623
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3011 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US95-01087-1

Query Match 11.0%; Score 13; DB 4; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 171
US-08-811-566-2
/ Sequence 2, Application US/08811566
/ Patent No. 6127116
/ GENERAL INFORMATION:
/ APPLICANT: Rice, Charles et al.
/ TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
/ TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: David A. Jackson, Esq.
/ STREET: 411 Hackensack Ave, Continental Plaza, 4th
/ STREET: Floor
/ CITY: Hackensack
```

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 03-MAR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1113-1-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 3012 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
US-08-811-566-2

Query Match 11.0%; Score 13; DB 2; Length 3012;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
DB 1664 GGVLAALAAAYCLS 1676

RESULT 172
US-09-034-756-2
Sequence 2, Application US/09034756
Patent No. 6392028
GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 04-May-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 3012 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
FRAGMENT TYPE: N-terminal
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-034-756-2

Query Match 11.0%; Score 13; DB 2; Length 3012;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
DB 1664 GGVLAALAAAYCLS 1676

RESULT 173
US-08-700-356-2
Sequence 2, Application US/08700356
Patent No. 5739002
GENERAL INFORMATION:
APPLICANT: DE FRANCESCO, Raffaele
APPLICANT: FAILLA, Cristina
APPLICANT: TOMBI, Licia
TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE
PROTEOLYTIC ACTIVITY OF THE NS3 HEPATITIS C VIRUS (HCV)
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 23-AUG-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: DE FRANCESCO-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-700-356-2

Query Match 10.2%; Score 12; DB 1; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 7 GGVLAALAAAYCL 18

RESULT 174
US-08-936-865-2

```
; Sequence 2, Application US/08936865
; Patent No. 5861297
; GENERAL INFORMATION:
; APPLICANT: Sardana, Vinod V
; APPLICANT: Blue, Jeffrey T
; TITLE OF INVENTION: DETERGENT-FREE HEPATITIS C PROTEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,865
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ayler, Sylvia A
; REGISTRATION NUMBER: 36,436
; REFERENCE/DOCKET NUMBER: 19691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-4909
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; LIBRARY: cDNA clone (See Seq. ID No. 5861297)
; CLONE: NS4A Protein
; US-08-936-865-2
;
; Query Match 10.2%; Score 12; DB 1; Length 54;
; Best Local Similarity 100.0%; Pred. No. 0.00027;
; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCL 29
; Db 7 GGVLAALAAAYCL 18
;
; RESULT 175
; US-09-198-723A-24
; Sequence 24, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 298-5056
; TELEFAX: (908) 298-5388
; INFORMATION FOR SEQ ID NO: 24:
;
; Query Match 10.2%; Score 12; DB 2; Length 54;
; Best Local Similarity 100.0%; Pred. No. 0.00027;
; Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 18 GGVLAALAAAYCL 29
; Db 7 GGVLAALAAAYCL 18
;
; RESULT 176
; US-09-684-881-24
; Sequence 24, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 298-5056
; TELEFAX: (908) 298-5388
; INFORMATION FOR SEQ ID NO: 24:
```


SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 24:
US-09-684-881-24

Query Match 10.2%; Score 12; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCL 29
| | | | | | | | | |
Db 7 GGVLAALAAYCL 18

RESULT 177

US-08-324-977-44
; Sequence 44, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142

INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-324-977-44

Query Match 10.2%; Score 12; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCL 29
| | | | | | | | | |
Db 49 GGVLAALAAYCL 60

RESULT 178

US-08-384-616-44
; Sequence 44, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-384-616-44

Query Match 10.2%; Score 12; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 49 GGVLAALAAAYCL 60

RESULT 179

US-08-304-686A-44
; Sequence 44, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-304-686A-44

Query Match 10.2%; Score 12; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 49 GGVLAALAAAYCL 60

RESULT 180

US-09-315-850-44
; Sequence 44, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357

; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-44

Query Match 10.2%; Score 12; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 49 GGVLAALAAAYCL 60

RESULT 181
US-08-971-036-2
; Sequence 2, Application US/08971036
; Patent No. 586684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-971-036-2

Query Match 10.2%; Score 12; DB 1; Length 675;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 628 GGVLAALAAAYCL 639

US-08-971-036-2
; Sequence 2, Application US/08971036
; Patent No. 586684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110

Query Match 10.2%; Score 12; DB 1; Length 675;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 628 GGVLAALAAAYCL 639

RESULT 182

US-09-096-570-2
; Sequence 2, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096.570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971.036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-096-570-2

Query Match 10.2%; Score 12; DB 2; Length 675;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 628 GGVLAALAAAYCL 639

RESULT 183

US-09-265-617B-2
; Sequence 2, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265.617B
; CURRENT FILING DATE: 1999-03-10

; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Vector and
; OTHER INFORMATION: Gene Fragments
US-09-265-617B-2

Query Match 10.2%; Score 12; DB 2; Length 675;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 628 GGVLAALAAAYCL 639
|||||

RESULT 184
US-09-263-933-4
; Sequence 4, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; EARLIER FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-4

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 943 GGVLAALAAAYCL 954
|||||

RESULT 185
US-09-263-933-11
; Sequence 11, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; EARLIER FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1692

; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-11

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 943 GGVLAALAAAYCL 954
|||||

RESULT 186
US-09-263-933-18
; Sequence 18, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; EARLIER FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-18

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 943 GGVLAALAAAYCL 954
|||||

RESULT 187
US-09-919-901-4
; Sequence 4, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-09-919-901-4

Query Match 10.2%; Score 12; DB 2; Length 1692;

Best Local Similarity 100.0%; Pred. No. 0.0065; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
Db 943 GGVLAALAAAYCL 954

RESULT 188
US-09-919-901-11
; Sequence 11, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-09-919-901-11

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
Db 943 GGVLAALAAAYCL 954

RESULT 189
US-09-919-901-18
; Sequence 18, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-09-919-901-18

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
|||||
Db 943 GGVLAALAAAYCL 954

RESULT 190
US-10-191-966-4
; Sequence 4, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-10-191-966-4

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
Db 943 GGVLAALAAAYCL 954

RESULT 191
US-10-191-966-11
; Sequence 11, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: :
US-10-191-966-11

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 18 GGVLAALAAAYCL 29
Db 943 GGVLAALAAAYCL 954

RESULT 192
US-10-191-966-18
; Sequence 18, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OP INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 1692
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
US-10-191-966-18

Query Match 10.2%; Score 12; DB 2; Length 1692;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 943 GGVLAALAAAYCL 954

RESULT 193
US-09-539-601-9
; Sequence 9, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-9

Query Match 10.2%; Score 12; DB 2; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 639 GGVLAALAAAYCL 650

RESULT 194
US-09-539-601-12
; Sequence 12, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-12

Query Match 10.2%; Score 12; DB 2; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 639 GGVLAALAAAYCL 650

RESULT 195
US-09-539-601-18
; Sequence 18, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-18

Query Match 10.2%; Score 12; DB 2; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 639 GGVLAALAAAYCL 650

RESULT 196
US-09-539-601-24
; Sequence 24, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 1985
; TYPE: PRT
US-09-539-601-24
```

; ORGANISM: Hepatitis C virus
US-09-539-601-24

Query Match 10.2%; Score 12; DB 2; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 639 GGVLAALAAAYCL 650

RESULT 197
US-09-539-601-30
; Sequence 30, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; PRIOR FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-30

Query Match 10.2%; Score 12; DB 2; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 639 GGVLAALAAAYCL 650

RESULT 198
US-10-259-275-42
; Sequence 42, Application US/10259275
; Patent No. 6921634
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0120
; CURRENT APPLICATION NUMBER: US/10/259,275
; CURRENT FILING DATE: 2003-01-13
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 42
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by the nucleotides 2119-8073 of
; OTHER INFORMATION: SEQ ID NO:41
US-10-259-275-42

Query Match 10.2%; Score 12; DB 2; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0076;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 639 GGVLAALAAAYCL 650

RESULT 199
US-08-324-977-12
; Sequence 12, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLealand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-324-977-12

Query Match 10.2%; Score 12; DB 1; Length 1985;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 200

US-08-384-616-12
; Sequence 12, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991

APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990

APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990

ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281

REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930

TELEFAX: (202) 887-0357
; TELEX: 440142

INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2013 amino acids

TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-384-616-12

Query Match 10.2%; Score 12; DB 1; Length 2013;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 201

US-08-904-686A-12
; Sequence 12, Application US/08904686A
; Patent No. 5998130

GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto

APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato

APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao

TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE

NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &

ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000

CITY: Washington
; STATE: D.C.

COUNTRY: U.S.A.
; ZIP: 20006

COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII

CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997

APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990

ATTORNEY/AGENT INFORMATION:
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281

REFERENCE/DOCKET NUMBER: 900703B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930

TELEFAX: (202) 887-0357
; TELEX: 440142

INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2013 amino acids

TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-904-686A-12

Query Match 10.2%; Score 12; DB 1; Length 2013;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||

Db 1664 GGVLAALAAAYCL 1675

RESULT 202
US-09-315-850-12
; Sequence 12, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armatrong, Westerman, Hattori, Mcleland &
; ADDRESS: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcleland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2013 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-12

Query Match 10.2%; Score 12; DB 2; Length 2013;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1664 GGVLAALAAAYCL 1675

RESULT 203
US-08-952-981A-2
; Sequence 2, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: TOMEI, Licia
; APPLICANT: BEHRENS, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
; TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
; FILE REFERENCE: IT0002P
; CURRENT APPLICATION NUMBER: US/08/952,981A
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: cDNA clone pCD (38-9.4)
US-08-952-981A-2

Query Match 10.2%; Score 12; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866

RESULT 204
US-09-539-601-6
; Sequence 6, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-6

Query Match 10.2%; Score 12; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 855 GGVLAALAAAYCL 866

RESULT 205
US-09-539-601-15
; Sequence 15, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System

```
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-15

Query Match      10.2%; Score 12; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      855 GGVLAALAAAYCL 866

RESULT 206
US-10-029-907-3
; Sequence 3, Application US/100299907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 882
; OTHER INFORMATION: Xaa is Lys or Arg
; NAME/KEY: VARIANT
; LOCATION: 1489
; OTHER INFORMATION: Xaa is Leu
US-10-029-907-3

Query Match      10.2%; Score 12; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      855 GGVLAALAAAYCL 866

RESULT 207
US-10-309-561a-3
; Sequence 3, Application US/10309561a
; Patent No. 6956117
; GENERAL INFORMATION:
; APPLICANT: Kukulj, George and Pause, Armin
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083-1-D1
; CURRENT APPLICATION NUMBER: US/10/309,561a
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US 10/029,907
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 10/257,857
```

```
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 882
; OTHER INFORMATION: Xaa is Lys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1489
; OTHER INFORMATION: Xaa is Leu
US-10-309-561a-3

Query Match      10.2%; Score 12; DB 2; Length 2201;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      855 GGVLAALAAAYCL 866

RESULT 208
US-09-263-933-2
; Sequence 2, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; EARLIER FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-2

Query Match      10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1035 GGVLAALAAAYCL 1046

RESULT 209
US-09-263-933-9
; Sequence 9, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; TITLE OF INVENTION: OF INHIBITORS OF THE HEPATITIS C VIRUS PROTEASE
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; CURRENT FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; EARLIER FILING DATE: 1998-08-05
```

```
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-9

Query Match      10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1035 GGVLAALAAAYCL 1046

RESULT 210
US-09-263-933-16
; Sequence 16, Application US/09263933
; Patent No. 6280940
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; EARLIER APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-263-933-16

Query Match      10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1035 GGVLAALAAAYCL 1046

RESULT 211
US-09-919-901-2
; Sequence 2, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-919-901-2

Query Match      10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1035 GGVLAALAAAYCL 1046

RESULT 212
US-09-919-901-9
; Sequence 9, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-919-901-9

Query Match      10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1035 GGVLAALAAAYCL 1046

RESULT 213
US-09-919-901-16
; Sequence 16, Application US/09919901
; Patent No. 6599738
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/09/919,901
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 09/263,933
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-919-901-16
```

US-09-919-901-16

Query Match 10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046

RESULT 214

US-10-191-966-2
; Sequence 2, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ;
US-10-191-966-2

Query Match 10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046

RESULT 215

US-10-191-966-9
; Sequence 9, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ;
US-10-191-966-9

Query Match 10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046

RESULT 216

US-10-191-966-16
; Sequence 16, Application US/10191966
; Patent No. 6790612
; GENERAL INFORMATION:
; APPLICANT: Potts, Karen E.
; APPLICANT: Jackson, Roberta L.
; APPLICANT: Patrick, Amy K.
; TITLE OF INVENTION: REPORTER GENE SYSTEM FOR USE IN CELL-BASED ASSESSMENT
; FILE REFERENCE: 0125-0005A
; CURRENT APPLICATION NUMBER: US/10/191,966
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US/09/263,933
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 09/129,611
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2307
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: ;
US-10-191-966-16

Query Match 10.2%; Score 12; DB 2; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0.0087;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1035 GGVLAALAAAYCL 1046

RESULT 217

US-08-324-977-32
; Sequence 32, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAWIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; FILE REFERENCE: 0125-0005A
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLan &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/099,706
FILING DATE: 30-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703D
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 2620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-324-977-32

Query Match 10.2%; Score 12; DB 1; Length 2620;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1275 GGVLAALAAAYCL 1286

RESULT 218
US-08-384-616-32
Sequence 32, Application US/08384616
Patent No. 5847101
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/384,616

FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/769,996
FILING DATE: 02-OCT-1991
APPLICATION NUMBER: JP 2-167466
FILING DATE: 25-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-230921
FILING DATE: 31-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 2-305605
FILING DATE: 09-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/635,451
FILING DATE: 28-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Stevens-Smith, Theresa M.
REGISTRATION NUMBER: 36,281
REFERENCE/DOCKET NUMBER: 900703B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 659-2930
TELEFAX: (202) 887-0357
TELEX: 440142
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 2620 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-384-616-32

Query Match 10.2%; Score 12; DB 1; Length 2620;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1275 GGVLAALAAAYCL 1286

RESULT 219
US-08-904-686A-32
Sequence 32, Application US/08904686A
Patent No. 598130
GENERAL INFORMATION:
APPLICANT: OKAYAMA, Hiroto
APPLICANT: FUKU, Isao
APPLICANT: MORI, Chisato
APPLICANT: TAKAMIZAWA, Akahisa
APPLICANT: YOSHIDA, Iwao
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
ADDRESS: Naughton
STREET: 1725 K St. N.W. Suite 1000
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/904,686A
FILING DATE: 01-AUG-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/324,977
FILING DATE: 18-OCT-1994

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-167466
;; FILING DATE: 25-JUN-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-230921
;; FILING DATE: 31-AUG-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-305605
;; FILING DATE: 09-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/099,706
;; FILING DATE: 30-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/769,996
;; FILING DATE: 02-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/635,451
;; FILING DATE: 28-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLeLland, Le-Nhung
;; REGISTRATION NUMBER: 31,541
;; REFERENCE/DOCKET NUMBER: 900703G
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 659-2930
;; TELEFAX: (202) 887-0357
;; INFORMATION FOR SEQ ID NO: 32:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2620 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-304-686A-32

Query Match 10.2%; Score 12; DB 1; Length 2620;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1275 GGVLAALAAAYCL 1286

RESULT 220
US-09-315-850-32
; Sequence 32, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESS: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,686

;; FILING DATE: 01-AUG-1997
;; APPLICATION NUMBER: US 08/324,977
;; FILING DATE: 18-OCT-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-167466
;; FILING DATE: 25-JUN-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-230921
;; FILING DATE: 31-AUG-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-305605
;; FILING DATE: 09-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/099,706
;; FILING DATE: 30-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/769,996
;; FILING DATE: 02-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/635,451
;; FILING DATE: 28-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLeLland, Le-Nhung
;; REGISTRATION NUMBER: 31,541
;; REFERENCE/DOCKET NUMBER: 900703G
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 659-2930
;; TELEFAX: (202) 887-0357
;; INFORMATION FOR SEQ ID NO: 32:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2620 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-315-850-32

Query Match 10.2%; Score 12; DB 2; Length 2620;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1275 GGVLAALAAAYCL 1286

RESULT 221
US-08-324-977-36
; Sequence 36, Application US/08324977
; Patent No. 5747339
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLland &
; ADDRESS: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,977

;; FILING DATE: 18-OCT-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-167466
;; FILING DATE: 25-JUN-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-230921
;; FILING DATE: 31-AUG-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-305605
;; FILING DATE: 09-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/059,706
;; FILING DATE: 30-JUL-1993
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/769,996
;; FILING DATE: 02-OCT-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/635,451
;; FILING DATE: 28-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Stevens-Smith, Theresa M.
;; REGISTRATION NUMBER: 36,281
;; REFERENCE/DOCKET NUMBER: 900703D
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 659-2930
;; TELEFAX: (202) 887-0357
;; TELEX: 440142
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2621 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-324-977-36

Query Match 10.2%; Score 12; DB 1; Length 2621;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCL 29
|||
Db 1275 GGVLAALAAYCL 1286

RESULT 222
US-08-384-616-36
; Sequence 36, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:

;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/769,996
;; FILING DATE: 02-OCT-1991
;; APPLICATION NUMBER: JP 2-167466
;; FILING DATE: 25-JUN-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-230921
;; FILING DATE: 31-AUG-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 2-305605
;; FILING DATE: 09-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/635,451
;; FILING DATE: 28-DEC-1990
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Stevens-Smith, Theresa M.
;; REGISTRATION NUMBER: 36,281
;; REFERENCE/DOCKET NUMBER: 900703B
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 659-2930
;; TELEFAX: (202) 887-0357
;; TELEX: 440142
;; INFORMATION FOR SEQ ID NO: 36:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2621 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-384-616-36

Query Match 10.2%; Score 12; DB 1; Length 2621;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCL 29
|||
Db 1275 GGVLAALAAYCL 1286

RESULT 223
US-08-904-686A-36
; Sequence 36, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcleland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-904-686A-36

Query Match 10.2%; Score 12; DB 1; Length 2621;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1275 GGVLAALAAAYCL 1286

RESULT 224
US-09-315-850-36
; Sequence 36, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, Mcleland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44MB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997

; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcleland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2621 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-36

Query Match 10.2%; Score 12; DB 2; Length 2621;
Best Local Similarity 100.0%; Pred. No. 0.0098;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1275 GGVLAALAAAYCL 1286

RESULT 225
US-10-259-275-40
; Sequence 40, Application US/10259275
; Patent No. 6921634
; GENERAL INFORMATION:
; APPLICANT: Lemon, Stanley M.
; APPLICANT: Yi, Minkyung
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE
; FILE REFERENCE: 265.0007 0120
; CURRENT APPLICATION NUMBER: US/10/259,275
; CURRENT FILING DATE: 2003-01-13
; PRIOR APPLICATION NUMBER: US 60/171,909
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/747,419
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/325,236
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US 60/338,123
; PRIOR FILING DATE: 2001-11-13
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 40
; LENGTH: 2985
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: amino acid sequence encoded by nucleotides 2077-11121 of SEQ ID
; OTHER INFORMATION: NO:39
US-10-259-275-40

Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 228
US-08-384-616-2
; Sequence 2, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-384-616-2

Query Match 10.2%; Score 12; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675
|||||

RESULT 229
US-08-384-616-14
; Sequence 14, Application US/08384616
; Patent No. 5847101
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLand &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/384,616
; FILING DATE:
; CLASSIFICATION: 424
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; NAME: Stevens-Smith, Theresa M.
; REGISTRATION NUMBER: 36,281
; REFERENCE/DOCKET NUMBER: 900703B
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; TELEX: 440142
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-384-616-14

Query Match 10.2%; Score 12; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29

Db 1664 GGVLAALAAAYCL 1675
|||||

RESULT 230
US-08-904-686A-2
; Sequence 2, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeeland, Le-Nhung
; REGISTRATION NUMBER: 900703G
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-904-686A-2

Query Match 10.2%; Score 12; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 231
US-08-904-686A-14
; Sequence 14, Application US/08904686A
; Patent No. 5998130
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeeland &
; ADDRESSEE: Naughton
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686A
; FILING DATE: 01-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeeland, Le-Nhung
; REGISTRATION NUMBER: 900703G
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-904-686A-14

Query Match 10.2%; Score 12; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

```
RESULT 232
US-09-014-416-3
; Sequence 3, Application US/09014416
; Patent No. 6153421
; GENERAL INFORMATION:
; APPLICANT: Yanagi, Masayuki
; APPLICANT: Bukh, Jens
; APPLICANT: Emerson, Susanne U.
; APPLICANT: Purcell, Robert H.
; TITLE OF INVENTION: CLONED GENOMES OF INFECTIOUS HEPATITIS C VIRUSES AND
; FILE REFERENCE: 20264276
; CURRENT APPLICATION NUMBER: US/09/014,416
; CURRENT FILING DATE: 1998-01-27
; EARLIER APPLICATION NUMBER: US 60/053,062
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-014-416-3

Query Match          10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1664 GGVLAALAAAYCL 1675

RESULT 233
US-09-850-2
; Sequence 2, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westernman, Hattori, McLeland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-2

Query Match          10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAALAAAYCL 29
Db      1664 GGVLAALAAAYCL 1675

RESULT 234
US-09-315-850-14
; Sequence 14, Application US/09315850
; Patent No. 6217872
; GENERAL INFORMATION:
; APPLICANT: OKAYAMA, Hiroto
; APPLICANT: FUKU, Isao
; APPLICANT: MORI, Chisato
; APPLICANT: TAKAMIZAWA, Akahisa
; APPLICANT: YOSHIDA, Iwao
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Armstrong, Westernman, Hattori, McLeland &
; STREET: 1725 K St. N.W. Suite 1000
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/315,850
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/904,686
; FILING DATE: 01-AUG-1997
; APPLICATION NUMBER: US 08/324,977
; FILING DATE: 18-OCT-1994
; PRIOR APPLICATION DATA:
```

; APPLICATION NUMBER: JP 2-167466
; FILING DATE: 25-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-230921
; FILING DATE: 31-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 2-305605
; FILING DATE: 09-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/099,706
; FILING DATE: 30-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/769,996
; FILING DATE: 02-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/635,451
; FILING DATE: 28-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: McLeland, Le-Nhung
; REGISTRATION NUMBER: 31,541
; REFERENCE/DOCKET NUMBER: 900703G
; TELEPHONE: (202) 659-2930
; TELEFAX: (202) 887-0357
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3010 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-315-850-14

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 235
US-09-539-601-3
; Sequence 3, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-3

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 236
US-09-539-601-21
; Sequence 21, Application US/09539601C

; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-21

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 237
US-09-539-601-27
; Sequence 27, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-27

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 238
US-09-539-601-33
; Sequence 33, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus

US-09-539-601-33

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 239

US-09-878-281A-101
Sequence 101, Application US/09878281A
Patent No. 6762024
GENERAL INFORMATION:
APPLICANT: Innogenetics N.V.
TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
TITLE OF INVENTION: and therapy
FILE REFERENCE: 35
CURRENT APPLICATION NUMBER: US/09/878,281A
CURRENT FILING DATE: 2001-06-12
NUMBER OF SEQ ID NOS: 284
SOFTWARE: PatentIn version 3.1
SEQ ID NO 101
LENGTH: 20
TYPE: PRT
ORGANISM: hepatitis C virus
US-09-878-281A-101

Query Match 9.3%; Score 11; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQRAATQQ 97
Db 10 LGLLQRAATQQ 20

RESULT 240

US-08-197-484-60
Sequence 60, Application US/08197484
Patent No. 6419331
GENERAL INFORMATION:
APPLICANT: VITIELLO, Maria A.
APPLICANT: CHESTNUT, Robert W.
APPLICANT: SETTE, Alessandro D.
APPLICANT: CELIS, Esteban
APPLICANT: GRAY, Howard
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
TITLE OF INVENTION: CTL IMMUNITY
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/197,484
FILING DATE: 16-FEB-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491

FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA: US 07/827,682
APPLICATION NUMBER: 29-JAN-1992
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA: US 07/749,568
APPLICATION NUMBER: 26-AUG-1991
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (206) 623-6793
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-197-484-60

Query Match 8.5%; Score 10; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAAAYCL 29
Db 1 VLAALAAAYCL 10

RESULT 241

US-08-197-484-139
Sequence 139, Application US/08197484
Patent No. 6419331
GENERAL INFORMATION:
APPLICANT: VITIELLO, Maria A.
APPLICANT: CHESTNUT, Robert W.
APPLICANT: SETTE, Alessandro D.
APPLICANT: CELIS, Esteban
APPLICANT: GRAY, Howard
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
TITLE OF INVENTION: CTL IMMUNITY
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/197,484
FILING DATE: 16-FEB-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,682
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991

ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4
TELEPHONE: (206) 467-9600
TELEFAX: (206) 623-6793
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-197-484-139

Query Match 8.5%; Score 10; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAAAYCL 29
Db 1 VLAALAAAYCL 10

RESULT 242
PCT-US95-02121-60
Sequence 60, Application PC/TUS9502121
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
TITLE OF INVENTION: CTL IMMUNITY
NUMBER OF SEQUENCES: 153
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02121
FILING DATE: 16-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,484
FILING DATE: 16-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,682
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4PC
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
PCT-US95-02121-60

Query Match 8.5%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAAAYCL 29
Db 1 VLAALAAAYCL 10

RESULT 243
PCT-US95-02121-139
Sequence 139, Application PC/TUS9502121
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
TITLE OF INVENTION: CTL IMMUNITY
NUMBER OF SEQUENCES: 153
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02121
FILING DATE: 16-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/197,484
FILING DATE: 16-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/935,811
FILING DATE: 26-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/874,491
FILING DATE: 27-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,682
FILING DATE: 29-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,568
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 14137-26-4PC
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
PCT-US95-02121-139

Query Match 8.5%; Score 10; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0048;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAAAYCL 29
Db 1 VLAALAAAYCL 10

RESULT 244
US-09-009-953-256
Sequence 256, Application US/09009953
Patent No. 6413517
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
TITLE OF INVENTION: Identification of Broadly

Reactive DR Restricted Epitopes

NUMBER OF SEQUENCES: 274
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/009,953
FILING DATE: 21-Jan-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/036,713
FILING DATE: 23-JAN-1997
APPLICATION NUMBER: US 60/037,432
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-011520US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 256:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 256:
US-09-009-953-256

Query Match 8.5%; Score 10; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.007;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAY 27
| | | | | | | | | |
Db 6 GGVLAALAAAY 15

RESULT 245
US-09-878-281A-62
; Sequence 62, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 128
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (13)..(13)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE

; LOCATION: (24)..(24)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (58)..(58)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (113)..(113)
; OTHER INFORMATION: "Xaa" is any amino acid
US-09-878-281A-62

Query Match 7.6%; Score 9; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACHSADLEV 9
| | | | | | | | | |
Db 2 ACHSADLEV 10

RESULT 246
US-09-878-281A-270
; Sequence 270, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 270
; LENGTH: 481
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-270

Query Match 7.6%; Score 9; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACHSADLEV 9
| | | | | | | | | |
Db 364 ACHSADLEV 372

RESULT 247
US-09-878-281A-198
; Sequence 198, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 198
; LENGTH: 484
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (263)..(263)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (337)..(337)


```
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (482)..(482)
; OTHER INFORMATION: "Xaa" is any amino acid
US-09-878-281A-198

Query Match          7.6%; Score 9; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 367 ACMSADLEV 375

RESULT 248
US-09-878-281A-200
; Sequence 1, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 200
; LENGTH: 484
; TYPE: PRT
; ORGANISM: hepatitis C virus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (228)..(228)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (303)..(303)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (353)..(353)
; OTHER INFORMATION: "Xaa" is any amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (378)..(378)
; OTHER INFORMATION: "Xaa" is any amino acid
US-09-878-281A-200

Query Match          7.6%; Score 9; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 367 ACMSADLEV 375

RESULT 249
US-08-700-356-1
; Sequence 1, Application US/08700356
; Patent No. 5739002
; GENERAL INFORMATION:
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: FAILIA, Cristina
; APPLICANT: TOMEI, Licia
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE PROTEOLYTIC ACTIVITY OF THE NS3 HEPATITIS C VIRUS (HCV)
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
```

```
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,356
; FILING DATE: 23-AUG-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: DE FRANCESCO-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-700-356-1

Query Match          7.6%; Score 9; DB 1; Length 631;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 621 ACMSADLEV 629

RESULT 250
US-08-936-865-1
; Sequence 1, Application US/08936865
; Patent No. 5861297
; GENERAL INFORMATION:
; APPLICANT: Sardana, Vinod V
; APPLICANT: Blue, Jeffrey T
; TITLE OF INVENTION: DETERGENT-FREE HEPATITIS C PROTEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,865
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ayler, Sylvia A
; REGISTRATION NUMBER: 36,436
; REFERENCE/DOCKET NUMBER: 19691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-4909
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 631 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C Virus
; STRAIN: NS3 Serine Protease Domain
; INDIVIDUAL ISOLATE: BK
; IMMEDIATE SOURCE:
; LIBRARY: described by Tomei et al. in 1993
; CLONE: cDNA clone pCD (38-9.4)
; POSITION IN GENOME:
; MAP POSITION: 1-180
US-08-936-865-1

Query Match          7.6%; Score 9; DB 1; Length 631;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACMSADLEV 9
Db      621 ACMSADLEV 629

RESULT 251
US-09-198-723A-23
; Sequence 23, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jave P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 632 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-198-723A-23

Query Match          7.6%; Score 9; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACMSADLEV 9
Db      622 ACMSADLEV 630

RESULT 252
US-09-684-881-23
; Sequence 23, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jave P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 632 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-684-881-23

Query Match          7.6%; Score 9; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACMSADLEV 9
Db      622 ACMSADLEV 630

RESULT 253
US-09-198-723A-60
; Sequence 60, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Weber, Patricia
```

APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
REFERENCE/DOCKET NUMBER: NS3 Protease and NS4A Cofactor Peptide
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-60

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 254
US-09-198-723A-63
Sequence 63, Application US/09198723A
Patent No. 6211338
GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:

NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-63
Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 255
US-09-198-723A-66
Sequence 66, Application US/09198723A
Patent No. 6211338
GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-66

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 256
US-09-198-723A-69
; Sequence 69, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-198-723A-69

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 257
US-09-198-723A-72
; Sequence 72, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-198-723A-72

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 258
US-09-684-881-60
; Sequence 60, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:

;
;
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-09-684-881-60

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 259

US-09-684-881-63
; Sequence 63, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NS3 Protease and NS4A Cofactor Peptide

NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723

FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.

REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388

INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 63:

US-09-684-881-63

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 260

US-09-684-881-66
; Sequence 66, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NS3 Protease and NS4A Cofactor Peptide

NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723

FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.

REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388

INFORMATION FOR SEQ ID NO: 66:
SEQUENCE CHARACTERISTICS:
LENGTH: 646 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 66:

US-09-684-881-66

Query Match 7.6%; Score 9; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 636 ACMSADLEV 644

RESULT 261

US-09-684-881-69
; Sequence 69, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NS3 Protease and NS4A Cofactor Peptide

NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA

```
;
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE DESCRIPTION: SEQ ID NO: 69:
;
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-684-881-69
;
; Query Match 7.6%; Score 9; DB 2; Length 646;
; Best Local Similarity 100.0%; Pred. No. 2.1;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 ACMSADLEV 9
; Db 636 ACMSADLEV 644
;
; RESULT 262
; US-09-684-881-72
; Sequence 72, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NS3 Protease and NS4A Cofactor Peptide
;
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 72:
; SEQUENCE DESCRIPTION: SEQ ID NO: 72:
;
; LENGTH: 646 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-09-684-881-72
;
; Query Match 7.6%; Score 9; DB 2; Length 646;
; Best Local Similarity 100.0%; Pred. No. 2.1;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 ACMSADLEV 9
; Db 636 ACMSADLEV 644
;
; RESULT 263
; US-09-543-376B-1
; Sequence 1, Application US/09543376B
; Patent No. 6524589
; GENERAL INFORMATION:
; APPLICANT: Reichert, Paul
; APPLICANT: Prorise, Winifred W.
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Yao, Nanhua
; APPLICANT: Weber, Patricia C.
; TITLE OF INVENTION: COMPOSITIONS OF HEPATITIS C VIRUS NS3/NS4A COMPLEX AND METHODS FOR
; FILE REFERENCE: IN01007
; CURRENT APPLICATION NUMBER: US/09/543,376B
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: HCV HIS-NS4A21-32-GSGS-NS3 (3-631)
; LOCATION: (1)..(665)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: HCVNS3 3-631
; LOCATION: (38)..(665)
; OTHER INFORMATION:
;
; US-09-543-376B-1
;
; Query Match 7.6%; Score 9; DB 2; Length 665;
; Best Local Similarity 100.0%; Pred. No. 2.1;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 ACMSADLEV 9
; Db 656 ACMSADLEV 664
;
; RESULT 264
; US-09-543-376B-2
; Sequence 2, Application US/09543376B
; Patent No. 6524589
; GENERAL INFORMATION:
; APPLICANT: Reichert, Paul
; APPLICANT: Prorise, Winifred W.
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Yao, Nanhua
; APPLICANT: Weber, Patricia C.
; TITLE OF INVENTION: COMPOSITIONS OF HEPATITIS C VIRUS NS3/NS4A COMPLEX AND METHODS FOR
; FILE REFERENCE: IN01007
; CURRENT APPLICATION NUMBER: US/09/543,376B
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: HCV HIS-NS4A21-32-GSGS-NS3 (3-631)
; LOCATION: (1)..(665)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: HCVNS3 3-631
; LOCATION: (38)..(665)
; OTHER INFORMATION:
;
; US-09-543-376B-1
;
; Query Match 7.6%; Score 9; DB 2; Length 665;
; Best Local Similarity 100.0%; Pred. No. 2.1;
; Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; Qy 1 ACMSADLEV 9
; Db 656 ACMSADLEV 664
;
; RESULT 264
; US-09-543-376B-2
; Sequence 2, Application US/09543376B
; Patent No. 6524589
; GENERAL INFORMATION:
; APPLICANT: Reichert, Paul
; APPLICANT: Prorise, Winifred W.
; APPLICANT: Taremi, Shahrar S.
; APPLICANT: Yao, Nanhua
; APPLICANT: Weber, Patricia C.
; TITLE OF INVENTION: COMPOSITIONS OF HEPATITIS C VIRUS NS3/NS4A COMPLEX AND METHODS FOR
; FILE REFERENCE: IN01007
; CURRENT APPLICATION NUMBER: US/09/543,376B
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: HCV HIS-NS4A21-32-GSGS-NS3 (3-631)
; LOCATION: (1)..(665)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: HCVNS3 3-631
; LOCATION: (38)..(665)
; OTHER INFORMATION:
;
; US-09-543-376B-1
```

```
; FILE REFERENCE: IN01007
; CURRENT APPLICATION NUMBER: US/09/543,376B
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-543-376B-2

Query Match          7.6%; Score 9; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACMSADLEV 9
Db      656 ACMSADLEV 664

RESULT 265
US-09-543-376B-3
; Sequence 3, Application US/09543376B
; Patent No. 6524589
; GENERAL INFORMATION:
; APPLICANT: Reichert, Paul
; APPLICANT: Prossie, Winifred W.
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Yao, Nanhua
; APPLICANT: Weber, Patricia C.
; TITLE OF INVENTION: COMPOSITIONS OF HEPATITIS C VIRUS NS3/NS4A COMPLEX AND METHODS FOR
; FILE REFERENCE: IN01007
; CURRENT APPLICATION NUMBER: US/09/543,376B
; CURRENT FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 665
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-543-376B-3

Query Match          7.6%; Score 9; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACMSADLEV 9
Db      656 ACMSADLEV 664

RESULT 266
US-09-198-723A-11
; Sequence 11, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-12

Query Match          7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACMSADLEV 9
Db      656 ACMSADLEV 664

RESULT 267
US-09-198-723A-12
; Sequence 12, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; MEDIUM TYPE: Floppy disk
COMPUTER READABLE FORM:
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-12
```

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACSADLEV 9
Db 656 ACSADLEV 664

RESULT 268
US-09-198-723A-13
; Sequence 13, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-198-723A-13

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACSADLEV 9
Db 656 ACSADLEV 664

RESULT 269
US-09-198-723A-14
; Sequence 14, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide

; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-198-723A-14

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACSADLEV 9
Db 656 ACSADLEV 664

RESULT 270
US-09-198-723A-15
; Sequence 15, Application US/09198723A
; Patent No. 6211338
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; APPLICANT: Taremi, Shahriar S.
; APPLICANT: Weber, Patricia
; APPLICANT: Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; TITLE OF INVENTION: NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800

TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-15

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 656 ACMSADLEV 664

RESULT 271

US-09-198-723A-16
Sequence 16, Application US/09198723A
Patent No. 6211338

GENERAL INFORMATION:
APPLICANT: Malcolim, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-16

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 656 ACMSADLEV 664

RESULT 272

US-09-198-723A-17
Sequence 17, Application US/09198723A
Patent No. 6211338

GENERAL INFORMATION:
APPLICANT: Malcolim, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-17

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 656 ACMSADLEV 664

RESULT 273

US-09-198-723A-18
Sequence 18, Application US/09198723A
Patent No. 6211338

GENERAL INFORMATION:
APPLICANT: Malcolim, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh

```
;
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723A
; FILING DATE: 24 NOV 1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-198-723A-18

Query Match          7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACMSADLEV 9
Db      656 ACMSADLEV 664

RESULT 274
US-09-684-881-11
; Sequence 11, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
;           Taremi, Shahrar S.
;           Yao, Nanhua
;           Weber, Patricia
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

```
;
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-684-881-11

Query Match          7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACMSADLEV 9
Db      656 ACMSADLEV 664

RESULT 275
US-09-684-881-12
; Sequence 12, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
;           Taremi, Shahrar S.
;           Yao, Nanhua
;           Weber, Patricia
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-684-881-12

Query Match          7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 ACMSADLEV 9
Db      656 ACMSADLEV 664

RESULT 276
US-09-684-881-13
; Sequence 13, Application US/09684881
; Patent No. 6653127
```

GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
Taremi, Shahrar S.
Weber, Patricia
Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NS3 Protease and NS4A Cofactor Peptide
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723
FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-684-881-13

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664

RESULT 277
US-09-684-881-14
Sequence 14, Application US/09684881
Patent No. 6653127
GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
Taremi, Shahrar S.
Weber, Patricia
Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NS3 Protease and NS4A Cofactor Peptide
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723
FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-684-881-13

COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723
FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-684-881-14

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664

RESULT 278
US-09-684-881-15
Sequence 15, Application US/09684881
Patent No. 6653127
GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
Taremi, Shahrar S.
Weber, Patricia
Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NS3 Protease and NS4A Cofactor Peptide
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723
FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-684-881-14

;
;
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 666 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 15:
US-09-684-881-15

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664

RESULT 279

US-09-684-881-16
; Sequence 16, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

;; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
;; NS3 Protease and NS4A Cofactor Peptide

;; NUMBER OF SEQUENCES: 123
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corp.
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07030

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Power Macintosh
;; OPERATING SYSTEM: 8.0.1
;; SOFTWARE: Microsoft Word 6.0.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/684,881
;; FILING DATE: 06-Oct-2000
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/198,723
;; FILING DATE: 24 NOV 1998

;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLaughlin, Jaye P.
;; REGISTRATION NUMBER: 41,211
;; REFERENCE/DOCKET NUMBER: JB0800
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908)298-5056
;; TELEFAX: (908)298-5388
;; INFORMATION FOR SEQ ID NO: 16:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 666 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-684-881-16

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664

RESULT 280

US-09-684-881-17
; Sequence 17, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

;; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
;; NS3 Protease and NS4A Cofactor Peptide

;; NUMBER OF SEQUENCES: 123
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corp.
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07030

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Power Macintosh
;; OPERATING SYSTEM: 8.0.1
;; SOFTWARE: Microsoft Word 6.0.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/684,881
;; FILING DATE: 06-Oct-2000
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/09/198,723
;; FILING DATE: 24 NOV 1998

;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLaughlin, Jaye P.
;; REGISTRATION NUMBER: 41,211
;; REFERENCE/DOCKET NUMBER: JB0800
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908)298-5056
;; TELEFAX: (908)298-5388
;; INFORMATION FOR SEQ ID NO: 17:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 666 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-684-881-17

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 656 ACMSADLEV 664

RESULT 281

US-09-684-881-18
; Sequence 18, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

;; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
;; NS3 Protease and NS4A Cofactor Peptide

;; NUMBER OF SEQUENCES: 123
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corp.
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth

STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/684,881
FILING DATE: 06-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723
FILING DATE: 24 NOV 1998
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 666 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Query Match 7.6%; Score 9; DB 2; Length 666;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 656 ACMSADLEV 664

RESULT 282
US-09-198-723A-19
Sequence 19, Application US/09198723A
Patent No. 6211338
GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 672 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-19

Query Match 7.6%; Score 9; DB 2; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 662 ACMSADLEV 670

RESULT 283
US-09-198-723A-20
Sequence 20, Application US/09198723A
Patent No. 6211338
GENERAL INFORMATION:
APPLICANT: Malcolm, Bruce
APPLICANT: Taremi, Shahrar S.
APPLICANT: Weber, Patricia
APPLICANT: Yao, Nanhua
TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
NUMBER OF SEQUENCES: 123
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07030
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Power Macintosh
OPERATING SYSTEM: 8.0.1
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/198,723A
FILING DATE: 24 NOV 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0800
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 672 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-198-723A-20

Query Match 7.6%; Score 9; DB 2; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
DB 662 ACMSADLEV 670

RESULT 284
US-09-684-881-19

; Sequence 19, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua
; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 672 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-684-881-19

Query Match 7.6%; Score 9; DB 2; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 662 ACMSADLEV 670

RESULT 285
US-09-684-881-20
; Sequence 20, Application US/09684881
; Patent No. 6653127
; GENERAL INFORMATION:
; APPLICANT: Malcolm, Bruce
; Taremi, Shahrar S.
; Weber, Patricia
; Yao, Nanhua

; TITLE OF INVENTION: Covalent Complexes of Hepatitis C Virus
; NS3 Protease and NS4A Cofactor Peptide
; NUMBER OF SEQUENCES: 123
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07030

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Power Macintosh
; OPERATING SYSTEM: 8.0.1
; SOFTWARE: Microsoft Word 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/684,881
; FILING DATE: 06-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/198,723
; FILING DATE: 24 NOV 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0800
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 672 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-684-881-20

Query Match 7.6%; Score 9; DB 2; Length 672;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 662 ACMSADLEV 670

RESULT 286
US-08-146-028-285
; Sequence 285, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 285:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-285

Query Match 6.8%; Score 8; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 2 DEMEECSQ 9

RESULT 287

US-08-146-028-291
; Sequence 291, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 291:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-291

Query Match 6.8%; Score 8; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 2 DEMEECSQ 9

RESULT 288

US-08-146-028-292
; Sequence 292, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 292:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-292

Query Match 6.8%; Score 8; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 1 DEMEECSQ 8

RESULT 289

US-08-146-028-298
; Sequence 298, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-298

Query Match 6.8%; Score 8; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 1 DEMEECSQ 8

RESULT 290

US-08-723-425A-285
; Sequence 285, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100

```
; INFORMATION FOR SEQ ID NO: 285:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 9 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-723-425A-285

Query Match          6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 2 DEMEECSQ 9

RESULT 291
US-08-723-425A-291
; Sequence 291, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
;   TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;   TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
;   TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 292:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 9 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-723-425A-292

Query Match          6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 293
US-08-723-425A-298
; Sequence 298, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
;   TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;   TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
;   TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```


;; APPLICATION NUMBER: US/08/723,425A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-13
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-816-4000
;; TELEFAX: 703-816-4100
;; INFORMATION FOR SEQ ID NO: 298:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-723-425A-298

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 1 DEMECSQ 8

RESULT 294

US-09-112-206-285
; Sequence 285, Application US/09112206
; Patent No. 6210903

; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/112,206
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/146,028

;; FILING DATE:
;; INFORMATION FOR SEQ ID NO: 285:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-112-206-285

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 2 DEMECSQ 9

RESULT 295

US-09-112-206-291
; Sequence 291, Application US/09112206
; Patent No. 6210903

; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/112,206
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/146,028

;; FILING DATE:
;; INFORMATION FOR SEQ ID NO: 291:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-112-206-291

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 2 DEMECSQ 9

RESULT 296

US-09-112-206-292
; Sequence 292, Application US/09112206
; Patent No. 6210903

; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/112,206
;; FILING DATE:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/146,028

;; FILING DATE:
;; INFORMATION FOR SEQ ID NO: 292:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-112-206-292

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | |
Db 1 DEMEECSQ 8

RESULT 297

US-09-112-206-298
; Sequence 298, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-298

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | |
Db 1 DEMEECSQ 8

RESULT 298

US-09-790-497A-275
; Sequence 275, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 275
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-275

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | |
Db 2 DEMEECSQ 9

RESULT 299

US-09-790-497A-276
; Sequence 276, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 276
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-276

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | |
Db 1 DEMEECSQ 8

RESULT 300

US-09-576-824A-275
; Sequence 275, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425

;; PRIOR FILING DATE: 1996-09-30
;; PRIOR APPLICATION NUMBER: 09/146,028
;; PRIOR FILING DATE: 1993-11-22
;; PRIOR APPLICATION NUMBER: PCT/EP93/00517
;; PRIOR FILING DATE: 1993-03-08
;; PRIOR APPLICATION NUMBER: EP 92400598.6
;; PRIOR FILING DATE: 1992-03-06
;; NUMBER OF SEQ ID NOS: 600
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 275
;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Hepatitis C virus
US-09-576-824A-275

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
DB 2 DEMECSQ 9

RESULT 301

US-09-576-824A-276
;; Sequence 276, Application US/09576824A
;; Patent No. 6667387
;; GENERAL INFORMATION:

;; APPLICANT: De Leva, Robert
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
;; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
;; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
;; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
;; TITLE OF INVENTION: EPITOPES. A PROCESS FOR PREPARING THEM AND COMPOSITIONS
;; TITLE OF INVENTION: CONTAINING THEM
;; FILE REFERENCE: 2752-11

;; CURRENT APPLICATION NUMBER: US/09/576,824A
;; CURRENT FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 08/723,425
;; PRIOR FILING DATE: 1996-09-30
;; PRIOR APPLICATION NUMBER: 09/146,028
;; PRIOR FILING DATE: 1993-11-22
;; PRIOR APPLICATION NUMBER: PCT/EP93/00517
;; PRIOR FILING DATE: 1993-03-08
;; PRIOR APPLICATION NUMBER: EP 92400598.6
;; PRIOR FILING DATE: 1992-03-06
;; NUMBER OF SEQ ID NOS: 600
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 276
;; LENGTH: 9
;; TYPE: PRT
;; ORGANISM: Hepatitis C virus
US-09-576-824A-276

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
DB 1 DEMECSQ 8

RESULT 302

US-09-680-497-285
;; Sequence 285, Application US/09680497
;; Patent No. 6709828
;; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED

;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 285:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-680-497-285

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
DB 2 DEMECSQ 9

RESULT 303

US-09-680-497-291
;; Sequence 291, Application US/09680497
;; Patent No. 6709828
;; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 291:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-680-497-291

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
DB 2 DEMECSQ 9

```
RESULT 304
US-09-680-497-292
; Sequence 292, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 292:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-292

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 305
US-09-680-497-298
; Sequence 298, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 298:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-298

Query Match 6.8%; Score 8; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 306
US-08-802-981-210
; Sequence 210, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-000300US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 210:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-802-981-210

Query Match 6.8%; Score 8; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 307
US-08-802-981-93
; Sequence 93, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
```

CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-0003000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product= "Aib"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /product= "Acp"
US-08-802-981-93

Query Match 6.8%; Score 8; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
| | | | | | | |
DB 6 DEMEECSQ 13

RESULT 308
US-09-747-287A-136
; Sequence 136, Application US/09747287A
; Patent No. 6893868
; GENERAL INFORMATION:
; APPLICANT: KOMORIYA, AKIRA
; APPLICANT: PACKARD, BEVERLY S.
; TITLE OF INVENTION: HOMO-DOUBLY LABELED COMPOSITIONS FOR THE DETECTION OF ENZYME
; FILE REFERENCE: 300-948600US
; CURRENT APPLICATION NUMBER: US/09/747,287A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/349,019
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US08/802,981
; PRIOR FILING DATE: 1997-02-20
; PRIOR APPLICATION NUMBER: PCT/US00/24882
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 136
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide.
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-747-287A-136

Query Match 6.8%; Score 8; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
| | | | | | | |
DB 6 DEMEECSQ 13

RESULT 309
US-09-394-019C-89
; Sequence 89, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903820US
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 89
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide substrate
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (3)..(3)
; OTHER INFORMATION: X is Aib
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)..(4)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)..(4)
; OTHER INFORMATION: X is epsilon aminocaproic acid
US-09-394-019C-89

Query Match 6.8%; Score 8; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
| | | | | | | |
DB 6 DEMEECSQ 13

RESULT 310
US-09-394-019C-334
; Sequence 334, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903820US
; FILE REFERENCE: 300-903820US

```
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 334
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide. Chemically synthesized protease substrate.
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is alpha-aminoisobutyric acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-394-019C-334

Query Match          6.8%; Score 8; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.74; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 6 DEMECSQ 13

RESULT 311
US-08-466-975A-10
; Sequence 10, Application US/08466975A
; Patent No. 5910404
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GEERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,975A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,671
; FILING DATE:
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
```

```
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466-975A-10

Query Match          6.8%; Score 8; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 13 DEMECSQ 20

RESULT 312
US-08-466-975A-11
; Sequence 11, Application US/08466975A
; Patent No. 5910404
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GEERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,975A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,671
; FILING DATE:
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
```

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-466-975A-11

Query Match 6.8%; Score 8; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 1 DEMEECSQ 8

RESULT 313

US-08-391-671A-10
; Sequence 10, Application US/08391671A
; Patent No. 5922532

; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671A
FILING DATE: 21-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 14-OCT-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-671A-10

Query Match 6.8%; Score 8; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 1 DEMEECSQ 8

RESULT 315

US-08-853-623D-26
; Sequence 26, Application US/08853623D
; Patent No. 5990276

; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Mui, Philip

QY 60 DEMEECSQ 67

Db 13 DEMEECSQ 20

RESULT 314

US-08-391-671A-11
; Sequence 11, Application US/08391671A
; Patent No. 5922532

; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671A
FILING DATE: 21-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100

INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-671A-11

Query Match 6.8%; Score 8; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 1 DEMEECSQ 8

APPLICANT: Weber, Patricia
TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,623D
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,470
FILING DATE: 10-MAY-1996
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: NS3/NS4A Cleavage site
US-08-853-623D-26

Query Match 6.8%; Score 8; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 1 CMSADLEV 8

RESULT 316
US-08-853-623D-27
Sequence 27, Application US/08853623D
Patent No. 5990276
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Mui, Philip
APPLICANT: Weber, Patricia
TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,623D
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,470
FILING DATE: 10-MAY-1996
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: NS4A/4B Cleavage Site
US-08-853-623D-27

Query Match 6.8%; Score 8; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 5 DEMEECSQ 12

RESULT 317
US-08-467-902A-10
Sequence 10, Application US/08467902A
Patent No. 6007982
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
APPLICANT: VAN HEUVERSUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,902A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671
FILING DATE: 14-OCT-1992
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-467-902A-10

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; 0; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 60 DEMECSQ 67
Db 13 DEMECSQ 20

RESULT 318

US-08-467-902A-11
Sequence 11, Application US/08467902A
Patent No. 6007982
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/467,902A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/391,671
FILING DATE:
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 9012424.1.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-467-902A-11

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; 0; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 60 DEMECSQ 67
Db 1 DEMECSQ 8

RESULT 319
US-09-208-966-47
Sequence 47, Application US/09208966
Patent No. 6221355
GENERAL INFORMATION:
APPLICANT: Dowdy, Steven F.
TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
FILE REFERENCE: 48881/1742
CURRENT APPLICATION NUMBER: US/09/208,966
CURRENT FILING DATE: 1998-12-10
EARLIER APPLICATION NUMBER: 60/082,402
EARLIER FILING DATE: 1998-04-20
EARLIER APPLICATION NUMBER: 60/069,012
EARLIER FILING DATE: 1997-12-10
NUMBER OF SEQ ID NOS: 57
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 47
LENGTH: 20
TYPE: PPT
ORGANISM: human
US-09-208-966-47

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; 0; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

QY 2 CMSADLEV 9
Db 1 CMSADLEV 8

RESULT 320

US-09-275-265-10
Sequence 10, Application US/09275265
Patent No. 6287761
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT J
APPLICANT: POLLET, DIRK
APPLICANT: MAERTENS, GEERT
APPLICANT: VAN HEUVERSWUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/275,265

```
;
; FILING DATE: 13-DEC-1991
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/391.671
; FILING DATE: 21-FEB-1995
; APPLICATION NUMBER: US 07/920.286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-275-265-10

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20

RESULT 321
US-09-275-265-11
; Sequence 11, Application US/09275265
; Patent No. 6287761
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; APPLICANT: POLLET, DIRK
; APPLICANT: MAERTENS, GEERT
; APPLICANT: VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; TITLE OF INVENTION: ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/275,265
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/391.671
; FILING DATE: 21-FEB-1995
; APPLICATION NUMBER: US 07/920.286
; FILING DATE: 14-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP91/02409
```

```
;
; FILING DATE: 13-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-275-265-11

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 322
US-08-850-328-8
; Sequence 8, Application US/08850328
; Patent No. 6379886
; GENERAL INFORMATION:
; APPLICANT: TAKAHAMA, Y.
; APPLICANT: SHIRAIISHI, J.
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS
; TITLE OF INVENTION: C VIRUS INFECTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,328
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mays, Thomas D
; REGISTRATION NUMBER: 34,524
; REFERENCE/DOCKET NUMBER: 32273-20004.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

US-08-850-328-8

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 323

US-09-941-611-10
; Sequence 10, Application US/09941611
; Patent No. 6576417
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,611
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: 1995-02-21
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-09-941-611-10

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20

RESULT 324

US-09-941-611-11

; Sequence 11, Application US/09941611
; Patent No. 6576417
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/941,611
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: 1995-02-21
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:

US-09-941-611-11

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 325

US-09-775-052A-47
; Sequence 47, Application US/09775052A
; Patent No. 6645501
; GENERAL INFORMATION:
; APPLICANT: Dowdy, Steven F.
; TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
; FILE REFERENCE: 48881/1742
; CURRENT APPLICATION NUMBER: US/09/775,052A
; CURRENT FILING DATE: 2001-12-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012

```
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 47
; LENGTH: 20
; TYPE: PRT
; ORGANISM: human
US-09-775-052A-47

Query Match          6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred.No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 1 CMSADLEV 8
|||||

RESULT 326
US-09-790-497A-50
; Sequence 50, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 50
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-50

Query Match          6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred.No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||

RESULT 327
US-09-790-497A-51
; Sequence 51, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 50
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-50

Query Match          6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred.No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20
|||||
```

```
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-51

Query Match          6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred.No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 7 DEMEECSQ 14
|||||

RESULT 328
US-09-790-497A-52
; Sequence 52, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 52
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-52

Query Match          6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred.No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8
|||||

RESULT 329
US-09-790-497A-121
```

; Sequence 121, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-121

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20

RESULT 330
US-09-576-824A-121
; Sequence 121, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-121

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20

RESULT 331
US-10-044-995-10
; Sequence 10, Application US/10044995
; Patent No. 6872520
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT J
; POLLET, DIRK
; MAERTENS, GEERT
; VAN HEUVERSWUN, HUGO
; TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
; ANTIBODIES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/044,995
; FILING DATE: 15-Jan-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/391,671
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 07/920,286
; FILING DATE: 14-OCT-1992
; APPLICATION NUMBER: WO PCT/EP91/02409
; FILING DATE: 13-DEC-1991
; APPLICATION NUMBER: EP 90124241.2
; FILING DATE: 14-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 7038164000
; TELEFAX: 7038164100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-044-995-10

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20

RESULT 332
US-10-044-995-11
; Sequence 11, Application US/10044995

Patent No. 6872520
GENERAL INFORMATION:
APPLICANT: DELLEYS, ROBERT J
POLLET, DIRK
MAERTENS, GEERT
VAN HEUVERSWUN, HUGO
TITLE OF INVENTION: SYNTHETIC ANTIGENS FOR THE DETECTION OF
ANTIBODIES TO HEPATITIS C VIRUS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/044,995
FILING DATE: 15-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/391,671
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/920,286
FILING DATE: 14-OCT-1992
APPLICATION NUMBER: WO PCT/EP91/02409
FILING DATE: 13-DEC-1991
APPLICATION NUMBER: EP 90124241.2
FILING DATE: 14-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 7038164000
TELEFAX: 7038164100
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-10-044-995-11

Query Match 6.8%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 1 DEMEECSQ 8

RESULT 333
PCT-US94-03407-11
Sequence 11, Application PC/TUS9405407
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: "NUCLEIC ACID TAGGED IMMUNOASSAY"
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG, P.C.
STREET: Suite 1200, 127 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: USA

ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05407
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/061,694
FILING DATE: 13-MAY-1993
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US94-05407-11

Query Match 6.8%; Score 8; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 13 DEMEECSQ 20

RESULT 334
US-08-146-028-50
Sequence 50, Application US/08146028
Patent No. 5891640
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,028
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-146-028-50

Query Match 6.8%; Score 8; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 14 DEMEECSQ 21

```
RESULT 335
US-08-146-028-51
; Sequence 51, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; NAME/KEY: Modified-site
; LOCATION: 22
; US-08-146-028-51

Query Match 6.8%; Score 8; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 8 DEMEECSQ 15

RESULT 336
US-08-146-028-52
; Sequence 52, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; NAME/KEY: Modified-site
; LOCATION: 22
; US-08-146-028-52

Query Match 6.8%; Score 8; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 8 DEMEECSQ 15
```

```
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-146-028-52

Query Match 6.8%; Score 8; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 2 DEMEECSQ 9

RESULT 337
US-08-146-028-121
; Sequence 121, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 1
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 22
US-08-146-028-121

Query Match 6.8%; Score 8; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 14 DEMEECSQ 21

RESULT 338
US-08-723-425A-50
; Sequence 50, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
```

NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-723-425A-50

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 14 DEMEECSQ 21

RESULT 339
US-08-723-425A-51
Sequence 51, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-723-425A-51

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 8 DEMEECSQ 15

RESULT 340
US-08-723-425A-52
Sequence 52, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHUYE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:

LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-723-425A-52

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
Db 2 DEMECSQ 9

RESULT 341

US-08-723-425A-121
Sequence 121, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Xaa is absent
LOCATION: 1
FEATURE:
NAME/KEY: Xaa is absent
LOCATION: 22

US-08-723-425A-121

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
Db 14 DEMECSQ 21

RESULT 342

US-09-112-206-50
Sequence 50, Application US/09112206
Patent No. 6210903
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/112,206
FILING DATE:
PRIOR APPLICATION NUMBER: US 08/146,028
FILING DATE:
INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-09-112-206-50

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
Db 14 DEMECSQ 21

RESULT 343

US-09-112-206-51
Sequence 51, Application US/09112206
Patent No. 6210903
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
Db 14 DEMECSQ 21

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
; US-09-112-206-51

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 8 DEMEECSQ 15

RESULT 344
US-09-112-206-52
; Sequence 52, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site

; LOCATION: 22
US-09-112-206-52

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 2 DEMEECSQ 9

RESULT 345
US-09-112-206-121
; Sequence 121, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: xaa is absent
; LOCATION: 1
; FEATURE:
; NAME/KEY: xaa is absent
; LOCATION: 22
; US-09-112-206-121

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | | | | |
Db 14 DEMEECSQ 21

RESULT 346
US-09-576-824A-50
; Sequence 50, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM

FILE REFERENCE: 2752-11
CURRENT APPLICATION NUMBER: US/09/576,824A
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 08/723,425
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: 09/146,028
PRIOR FILING DATE: 1993-11-22
PRIOR APPLICATION NUMBER: PCT/EP93/00517
PRIOR FILING DATE: 1993-03-08
PRIOR APPLICATION NUMBER: EP 92400598.6
PRIOR FILING DATE: 1992-03-06
NUMBER OF SEQ ID NOS: 600
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 50
LENGTH: 22
TYPE: PRT
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: modified site
NAME/KEY: VARIANT
LOCATION: (22)
OTHER INFORMATION: modified site
US-09-576-824A-50

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 14 DEMEECSQ 21

RESULT 347
US-09-576-824A-51
Sequence 51, Application US/09576824A
Patent No. 6667387
GENERAL INFORMATION:
APPLICANT: De Leys, Robert
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
TITLE OF INVENTION: CONTAINING THEM
FILE REFERENCE: 2752-11
CURRENT APPLICATION NUMBER: US/09/576,824A
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 08/723,425
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: 09/146,028
PRIOR FILING DATE: 1993-11-22
PRIOR APPLICATION NUMBER: PCT/EP93/00517
PRIOR FILING DATE: 1993-03-08
PRIOR APPLICATION NUMBER: EP 92400598.6
PRIOR FILING DATE: 1992-03-06
NUMBER OF SEQ ID NOS: 600
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 51
LENGTH: 22
TYPE: PRT
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: modified site
NAME/KEY: VARIANT
LOCATION: (22)
OTHER INFORMATION: modified site
US-09-576-824A-51

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 8 DEMEECSQ 15

RESULT 348
US-09-576-824A-52
Sequence 52, Application US/09576824A
Patent No. 6667387
GENERAL INFORMATION:
APPLICANT: De Leys, Robert
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
TITLE OF INVENTION: CONTAINING THEM
FILE REFERENCE: 2752-11
CURRENT APPLICATION NUMBER: US/09/576,824A
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 08/723,425
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: 09/146,028
PRIOR FILING DATE: 1993-11-22
PRIOR APPLICATION NUMBER: PCT/EP93/00517
PRIOR FILING DATE: 1993-03-08
PRIOR APPLICATION NUMBER: EP 92400598.6
PRIOR FILING DATE: 1992-03-06
NUMBER OF SEQ ID NOS: 600
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 52
LENGTH: 22
TYPE: PRT
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: modified site
NAME/KEY: VARIANT
LOCATION: (22)
OTHER INFORMATION: modified site
US-09-576-824A-52

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 2 DEMEECSQ 9

RESULT 349
US-09-680-497-50
Sequence 50, Application US/09680497
Patent No. 6709828
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 50:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: HCV
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 22
;; US-09-680-497-50

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 60 DEMEECSQ 67
Db 14 DEMEECSQ 21

RESULT 350
US-09-680-497-51
; Sequence 51, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
; US-09-680-497-51

Query Match 6.8%; Score 8; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 0.85; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 60 DEMEECSQ 67
Db 8 DEMEECSQ 15

RESULT 351
US-09-680-497-52
; Sequence 52, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
; US-09-680-497-52

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0;

Qy 60 DEMEECSQ 67
Db 2 DEMEECSQ 9

RESULT 352
US-09-680-497-121
; Sequence 121, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 121:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: HCV
;; FEATURE:
;; NAME/KEY: Xaa is absent
;; LOCATION: 1
;; FEATURE:
;; NAME/KEY: Xaa is absent
;; LOCATION: 22
US-09-680-497-121

Query Match 6.8%; Score 8; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMERCSQ 67
|||||||
DB 14 DEMERCSQ 21

RESULT 353
PCT-US92-07865-1
;; Sequence 1, Application PC/TUS9207865
;; GENERAL INFORMATION:
;; APPLICANT: Dreesman, Gordon R.
;; APPLICANT: Burk, Kenneth H.
;; APPLICANT: Pauletti, Daniel
;; TITLE OF INVENTION: Peptide-Based Hepatitis C Virus
;; TITLE OF INVENTION: Immunoassays
;; NUMBER OF SEQUENCES: 25
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Law Offices of Peter Dehlinger
;; STREET: 350 Cambridge Avenue, Suite 300
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94306
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/07865
;; FILING DATE: 19920916
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; FILING DATE: 16-SEP-1991
;; FILING DATE: 12-FEB-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fabian, Gary R.
;; REGISTRATION NUMBER: 33,875
;; REFERENCE/DOCKET NUMBER: 1600-0086.41
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-0880
;; TELEFAX: (415) 324-0960
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 24 amino acids
;; TYPE: AMINO ACID

;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
;; ORIGINAL SOURCE:
;; ORGANISM: Hepatitis C Virus
;; INDIVIDUAL ISOLATE: Dp1, amino acids 1694 to 1717 of HCV
;; INDIVIDUAL ISOLATE: polyprotein
PCT-US92-07865-1

Query Match 6.8%; Score 8; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.92;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMERCSQ 67
|||||||
DB 13 DEMERCSQ 20

RESULT 354
US-09-929-955-25
;; Sequence 25, Application US/09929955
;; Patent No. 6858590
;; GENERAL INFORMATION:
;; APPLICANT: Matti Sallberg
;; APPLICANT: Catharina Hultgren
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: TRIPEP.23AUS2
;; CURRENT APPLICATION NUMBER: US/09/929,955
;; CURRENT FILING DATE: 2001-08-15
;; PRIOR APPLICATION NUMBER: 09/705,547
;; PRIOR FILING DATE: 2000-11-03
;; PRIOR APPLICATION NUMBER: 60/229,175
;; PRIOR FILING DATE: 2000-08-29
;; PRIOR APPLICATION NUMBER: 60/225,767
;; PRIOR FILING DATE: 2000-08-17
;; NUMBER OF SEQ ID NOS: 49
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 25
;; LENGTH: 25
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Hepatitis C virus NS3/4A peptide
US-09-929-955-25

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
|||||||
DB 6 CMSADLEV 13

RESULT 355
US-09-929-955-26
;; Sequence 26, Application US/09929955
;; Patent No. 6858590
;; GENERAL INFORMATION:
;; APPLICANT: Matti Sallberg
;; APPLICANT: Catharina Hultgren
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: TRIPEP.23AUS2
;; CURRENT APPLICATION NUMBER: US/09/929,955
;; CURRENT FILING DATE: 2001-08-15
;; PRIOR APPLICATION NUMBER: 09/705,547
;; PRIOR FILING DATE: 2000-11-03
;; PRIOR APPLICATION NUMBER: 60/229,175
;; PRIOR FILING DATE: 2000-08-29

; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-26

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 356

US-09-929-955-27
; Sequence 27, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-27

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 357

US-09-929-955-33
; Sequence 33, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175

; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-33

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 358

US-09-929-955-34
; Sequence 34, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-34

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 359

US-09-929-955-35
; Sequence 35, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03

; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-35

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
| | | | |
Db 6 CMSADLEV 13

RESULT 360
US-09-929-955-36
; Sequence 36, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-36

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
| | | | |
Db 6 CMSADLEV 13

RESULT 361
US-09-929-955-37
; Sequence 37, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547

; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-37

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
| | | | |
Db 6 CMSADLEV 13

RESULT 362
US-09-929-955-38
; Sequence 38, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-38

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
| | | | |
Db 6 CMSADLEV 13

RESULT 363
US-09-930-591-14
; Sequence 14, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; TITLE OF INVENTION: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767

```
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS3/4A peptide
US-09-930-591-14

Query Match          6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 364
US-09-930-591-16
; Sequence 16, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-16

Query Match          6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 365
US-09-930-591-17
; Sequence 17, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
```

```
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-17

Query Match          6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 366
US-09-930-591-18
; Sequence 18, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-18

Query Match          6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 367
US-09-930-591-19
; Sequence 19, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
```


; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-19

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 368
US-09-930-591-20
; Sequence 20, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-20

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 369
US-09-930-591-21
; Sequence 21, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-21

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 370
US-09-930-591-22
; Sequence 22, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-22

Query Match 6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 6 CMSADLEV 13

RESULT 371
US-09-930-591-23
; Sequence 23, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547

```
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-23

Query Match          6.8%; Score 8; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.95; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Qy      2 CMSADLEV 9
Db      6 CMSADLEV 13

RESULT 372
US-07-946-054-7
; Sequence 7, Application US/07946054
; Patent No. 5582968
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Hosein, Barbara H
; TITLE OF INVENTION: No. 5582968el Branched Hybrid and Cluster
; TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-A,
; TITLE OF INVENTION: No. 5582968-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: United Biomedical Inc.
; STREET: 25 Davids Dr.
; CITY: Hauppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946.054
; FILING DATE: 15-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, M. Lisa
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 2000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-273-2828
; TELEFAX: 516-273-1717
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-946-054-7

Query Match          6.8%; Score 8; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEECSQ 67
Db      18 DEMEECSQ 25

RESULT 373
US-08-336-553A-17
; Sequence 17, Application US/08336553A
; Patent No. 6054264
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; APPLICANT: KUO, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; TITLE OF INVENTION: REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-336-553A-17

Query Match          6.8%; Score 8; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEECSQ 67
Db      18 DEMEECSQ 25

RESULT 374
US-08-439-157-17
; Sequence 17, Application US/08439157
; Patent No. 6416944
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; APPLICANT: KUO, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; TITLE OF INVENTION: REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,157
FILING DATE: 11-May-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553A
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-08-439-157-17
Query Match 6.8%; Score 8; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 18 DEMEECSQ 25
RESULT 375
US-09-437-895-17
Sequence 17, Application US/09437895
Patent No. 6416946
GENERAL INFORMATION:
APPLICANT: CHIEN, DAVID Y.
KUO, GEORGE
TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
REAGENTS FOR USE THEREIN
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/437,895
FILING DATE: 09-No. 6416946-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-437-895-17
Query Match 6.8%; Score 8; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 18 DEMEECSQ 25
RESULT 376
PCT-US93-08638-7
Sequence 7, Application PC/TUS9308638
GENERAL INFORMATION:
APPLICANT: United Biomedical Inc.
TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
Effective in Diagnosing and Detecting Non-A,
TITLE OF INVENTION: Non-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: UNITED BIOMEDICAL INC.
STREET: 25 Davids Drive
CITY: Hauppauge
STATE: New York
COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08638
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: M. Lisa Wilson
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 9055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
TELEX:
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08638-7
Query Match 6.8%; Score 8; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECSQ 67
Db 18 DEMEECSQ 25
RESULT 377
US-07-666-719-11
Sequence 11, Application US/07666719

```
; Patent No. 5247067
; GENERAL INFORMATION:
; APPLICANT: ARIMA, Terukatsu
; APPLICANT: YAMADA, Kyoko
; APPLICANT: HATANAKA, Tadashi
; APPLICANT: NABA, Toshihiko
; APPLICANT: TSUI, Masao
; TITLE OF INVENTION: PEPTIDE AND ITS USE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/666,719
; FILING DATE: 19910422
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5247067man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 363-264-0X
; TELEPHONE: (703)521-5940
; TELEFAX: (703)486-2347
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-666-719-11

Query Match 6.8%; Score 8; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.2; Indels 0;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 14 DEMEECSQ 21

RESULT 378
US-07-666-719-5
; Sequence 5, Application US/07666719
; Patent No. 5247067
; GENERAL INFORMATION:
; APPLICANT: ARIMA, Terukatsu
; APPLICANT: YAMADA, Kyoko
; APPLICANT: HATANAKA, Tadashi
; APPLICANT: NABA, Toshihiko
; APPLICANT: TSUI, Masao
; TITLE OF INVENTION: PEPTIDE AND ITS USE
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: US
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/666,719
; FILING DATE: 19910422
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5247067man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 363-264-0X
; TELEPHONE: (703)521-5940
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 39 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-666-719-5

Query Match 6.8%; Score 8; DB 1; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
Db 22 DEMEECSQ 29

RESULT 379
US-08-444-818-771
; Sequence 771, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 771:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 41 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-444-818-771
```

Query Match 6.8%; Score 8; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 60 DEMERCSQ 67
DB 13 DEMERCSQ 20

RESULT 380

US-08-905-054B-13
; Sequence 13, Application US/08905054B
; Patent No. 6596476
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Lesniewski, Richard R.
; APPLICANT: Leung, Tat K.
; TITLE OF INVENTION: HEPATITIS C ASSAY
; FILE REFERENCE: 4767-US-C7
; CURRENT APPLICATION NUMBER: US/08/905,054B
; CURRENT FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/707,355
; PRIOR FILING DATE: 1996-09-04
; PRIOR APPLICATION NUMBER: US 08/507,740
; PRIOR FILING DATE: 1995-07-26
; PRIOR APPLICATION NUMBER: US 08/373,920
; PRIOR FILING DATE: 1995-01-17
; PRIOR APPLICATION NUMBER: US 08/183,207
; PRIOR FILING DATE: 1994-01-18
; PRIOR APPLICATION NUMBER: US 07/760,292
; PRIOR FILING DATE: 1991-09-16
; PRIOR APPLICATION NUMBER: US 07/610,180
; PRIOR FILING DATE: 1990-11-07
; PRIOR APPLICATION NUMBER: US 07/456,162
; PRIOR FILING DATE: 1989-12-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-08-905-054B-13

Query Match 6.8%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 60 DEMERCSQ 67
DB 13 DEMERCSQ 20

RESULT 381

PCT-US92-07813-13
; Sequence 13, Application PC/TUS9207813
; GENERAL INFORMATION:
; APPLICANT: LESNIEWSKI, RICHARD R.
; APPLICANT: LEUNG, TAT K.
; TITLE OF INVENTION: HEPATITIS C ASSAY
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: ABBOTT LABORATORIES CHAD377/AP6D
; CITY: ABBOTT PARK
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07813
; FILING DATE: 19920916
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKIP, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4767.P3.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US92-07813-13

Query Match 6.8%; Score 8; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0;

QY 60 DEMERCSQ 67
DB 13 DEMERCSQ 20

RESULT 382

US-07-946-054-6
; Sequence 6, Application US/07946054
; Patent No. 5582968
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Hosein, Barbara H
; TITLE OF INVENTION: No. 5582968el Branched Hybrid and Cluster
; TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-B Hepatitis
; TITLE OF INVENTION: No. 5582968-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: United Biomedical Inc.
; STREET: 25 Davids Dr.
; CITY: Hauppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,054
; FILING DATE: 15-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, M. Lisa
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 2000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-273-2828
; TELEFAX: 516-273-1717
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-946-054-6

Query Match 6.8%; Score 8; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;

```
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 18 DEMECSQ 25

RESULT 383
US-08-083-947-22
; Sequence 22, Application US/08083947
; Patent No. 5639594
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Hosein, Barbara
; TITLE OF INVENTION: NO. 5639594el Linear and Branched Peptides Effective
; TITLE OF INVENTION: In Diagnosing and Detecting No. 5639594-A, NO. 5639594-B Hepat
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: M. Lisa Wilson
; STREET: 25 Davids Drive
; CITY: Hauppauge
; STATE: NY
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,947
; FILING DATE: 19930628
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 946,054
; FILING DATE: 15-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, M. Lisa
; REGISTRATION NUMBER: 34045
; REFERENCE/DOCKET NUMBER: 2000Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)273-2828
; TELEFAX: (516)273-1717
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-083-947-22

Query Match 6.8%; Score 8; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 18 DEMECSQ 25

RESULT 384
US-08-530-550-2
; Sequence 2, Application US/08530550
; Patent No. 5736321
; GENERAL INFORMATION:
; APPLICANT: Hosein, Barbara
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Peptides Effective for Diagnosis and
; TITLE OF INVENTION: Detection of Hepatitis c Infection
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: M. Lisa Wilson
; STREET: 25 Davids Drive
```

```
; CITY: Hauppauge
; STATE: NY
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/530,550
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, M. Lisa
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 2000Z
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516)273-2828
; TELEFAX: (516)273-1717
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-530-550-2

Query Match 6.8%; Score 8; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECSQ 67
Db 18 DEMECSQ 25

RESULT 385
US-08-262-037-21
; Sequence 21, Application US/08262037
; Patent No. 5747239
; GENERAL INFORMATION:
; APPLICANT: Chang Yi Wang and Barbara Hosein
; TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
; TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
; NUMBER OF SEQUENCES: 136
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVE.
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/262,037
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/719,819
; FILING DATE: 24-June-1991
; APPLICATION NUMBER: 07/667,275
; FILING DATE: 11-Mar-1991
; APPLICATION NUMBER: 07/651,735
; FILING DATE: 07-Feb-1991
; APPLICATION NUMBER: 07/558,799
; FILING DATE: 26-July-1990
; APPLICATION NUMBER: 07/510,153
```

```

; FILING DATE: 16-APRIL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C. H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4043 US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: Amino acid
; STRANDEDNESS:
; TOPOLOGY: Unknown
; US-08-262-037-21

Query Match 6.8%; Score 8; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 18 DEMEECSQ 25

RESULT 386
PCT-US93-08638-6
; Sequence 6, Application PC/TUS9308638
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc.
; TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
; TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
; TITLE OF INVENTION: Non-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: UNITED BIOMEDICAL INC.
; STREET: 25 Davids Drive
; CITY: Huppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08638
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,947
; FILING DATE: 28-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIN, MARIA C.H.
; REGISTRATION NUMBER: 29323
; REFERENCE/DOCKET NUMBER: 1151-4101PC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08638-6

Query Match 6.8%; Score 8; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 18 DEMEECSQ 25

RESULT 387
PCT-US94-07088-22
; Sequence 22, Application PC/TUS9407088
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Hosein, Barbara
; TITLE OF INVENTION: Novel Linear And Branched
; TITLE OF INVENTION: Peptides Effective In
; TITLE OF INVENTION: Diagnosing And Detecting
; TITLE OF INVENTION: Non-A, Non-B Hepatitis
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07088
; FILING DATE: 22-JUNE-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,947
; FILING DATE: 28-JUNE-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LIN, MARIA C.H.
; REGISTRATION NUMBER: 29323
; REFERENCE/DOCKET NUMBER: 1151-4101PC1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US94-07088-22

Query Match 6.8%; Score 8; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
DB 18 DEMEECSQ 25

RESULT 388
PCT-US95-13660-2
; Sequence 2, Application PC/TUS9513660
; GENERAL INFORMATION:
; APPLICANT: Hosein, Barbara
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: Peptides Effective for
; TITLE OF INVENTION: Diagnosis and Detection of Hepatitis C Infection
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; US-09-638-693a-36_copy_16_133.Orig.ra1
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13660
; FILING DATE: 23 October 1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/333,573
; FILING DATE: 01 November 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4118PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)758-4800
; TELEFAX: (212)751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US95-13660-2

Query Match 6.8%; Score 8; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
Db 18 DEMEECSQ 25

RESULT 389
US-07-946-054-11
; Sequence 11, Application US/07946054
; Patent No. 5582968
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Hosein, Barbara H
; TITLE OF INVENTION: No. 5582968a1 Branched Hybrid and Cluster
; TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-A,
; TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: United Biomedical Inc.
; STREET: 25 Davids Dr.
; CITY: Hauppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,054
; FILING DATE: 15-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, M. Lisa
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 2000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-273-2828
; TELEFAX: 516-273-1717
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-946-054-11

Query Match 6.8%; Score 8; DB 1; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
Db 21 DEMEECSQ 28

RESULT 390
PCT-US93-08638-11
; Sequence 11, Application PC/TUS9308638
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc.
; TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
; TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
; TITLE OF INVENTION: Non-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: UNITED BIOMEDICAL INC.
; STREET: 25 Davids Drive
; CITY: Hauppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08638
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: M. Lisa Wilson
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 9055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-273-2828
; TELEFAX: 516-273-1717
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08638-11

Query Match 6.8%; Score 8; DB 4; Length 50;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECSQ 67
Db 21 DEMEECSQ 28

RESULT 391
US-08-444-818-2
; Sequence 2, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
```


NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-2

Query Match 6.8%; Score 8; DB 2; Length 51;
Best Local Similarity 100.0%; Pred.No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||
Db 22 DEMEECSQ 29

RESULT 392
US-08-905-054B-11
Sequence 11, Application US/08905054B
Patent No. 6596476
GENERAL INFORMATION:
APPLICANT: Abbott Laboratories
ADDRESSEE: Lesniewski, Richard R.
STREET: Leung, Tat K.
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-818-2

Query Match 6.8%; Score 8; DB 2; Length 51;
Best Local Similarity 100.0%; Pred.No. 1.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||
Db 22 DEMEECSQ 29

LENGTH: 67
TYPE: PRT
ORGANISM: Hepatitis C Virus
US-08-905-054B-11

Query Match 6.8%; Score 8; DB 2; Length 67;
Best Local Similarity 100.0%; Pred.No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||
Db 23 DEMEECSQ 30

RESULT 393
PCT-US92-07813-11
Sequence 11, Application PC/TUS9207813
GENERAL INFORMATION:
APPLICANT: LESNIEWSKI, RICHARD R.
ADDRESSEE: LEUNG, TAT K.
TITLE OF INVENTION: HEPATITIS C ASSAY
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES CHAD377/AP60
STREET: ONE ABBOTT PARK ROAD
CITY: ABBOTT PARK
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/07813
FILING DATE: 19920916
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FOREMSKIP, PRISCILLA E.
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4767.P3.03
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US92-07813-11

Query Match 6.8%; Score 8; DB 4; Length 67;
Best Local Similarity 100.0%; Pred.No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||
Db 23 DEMEECSQ 30

RESULT 394
US-08-444-818-10
Sequence 10, Application US/08444818
Patent No. 6150087
GENERAL INFORMATION:
APPLICANT: Chien, David Y.
ADDRESSEE: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:

```
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-444-818-10

Query Match      6.8%; Score 8; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      60 DEMEECSQ 67
Db      18 DEMEECSQ 25

RESULT 395
US-08-905-054B-12
; Sequence 12, Application US/08905054B
; Patent No. 6596476
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Lesniewski, Richard R.
; TITLE OF INVENTION: HEPATITIS C ASSAY
; CURRENT APPLICATION NUMBER: US/08/905,054B
; CURRENT FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/707,355
; PRIOR FILING DATE: 1996-09-04
; PRIOR APPLICATION NUMBER: US 08/507,740
; PRIOR FILING DATE: 1995-07-26
; PRIOR APPLICATION NUMBER: US 08/373,920
; PRIOR FILING DATE: 1995-01-17
; PRIOR APPLICATION NUMBER: US 08/183,207
; PRIOR FILING DATE: 1994-01-18
; PRIOR APPLICATION NUMBER: US 07/760,292
; PRIOR FILING DATE: 1991-09-16
; PRIOR APPLICATION NUMBER: US 07/610,180
; PRIOR FILING DATE: 1990-11-07
; PRIOR APPLICATION NUMBER: US 07/456,162
; PRIOR FILING DATE: 1989-12-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 117
; TYPE: PRT
```

```
; ORGANISM: Hepatitis C Virus
; US-08-905-054B-12

Query Match      6.8%; Score 8; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 4;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      60 DEMEECSQ 67
Db      18 DEMEECSQ 25

RESULT 396
PCT-US92-07813-12
; Sequence 12, Application PC/TUS9207813
; GENERAL INFORMATION:
; APPLICANT: LESNIEWSKI, RICHARD R.
; APPLICANT: LEUNG, TAT K.
; TITLE OF INVENTION: HEPATITIS C ASSAY
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES CHAD377/AP6D
; STREET: ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: ILLINOIS
; COUNTRY: U.S.A.
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/07813
; FILING DATE: 19920916
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKIP, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4767.P3.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US92-07813-12

Query Match      6.8%; Score 8; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 4;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Qy      60 DEMEECSQ 67
Db      18 DEMEECSQ 25

RESULT 397
US-08-444-818-8
; Sequence 8, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
```

; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-444-818-8

Query Match 6.8%; Score 8; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
|||||||

Db 29 DEMEECSQ 36

RESULT 398
US-08-833-678A-2
; Sequence 2, Application US/08833678A
; Patent No. 5989905
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI-LIM
; APPLICANT: HAN, JANG
; APPLICANT: CHOE, JOONHO
; TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: Intellectual Property - R440, P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,678A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/529,169
; FILING DATE: 15-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.

; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0100.005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3274
; TELEFAX: (510) 655-3542
; TELEX: n/a
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-833-678A-2

Query Match 6.8%; Score 8; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||||

Db 456 CMSADLEV 463

RESULT 399
US-08-529-169A-2
; Sequence 2, Application US/08529169A
; Patent No. 6194140
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI-LIM
; APPLICANT: HAN, JANG
; APPLICANT: CHOE, JOONHO
; TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: Intellectual Property - R440, P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/529,169A
; FILING DATE: 15-SEP-1995
; CLASSIFICATION: 4325
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0100.005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3274
; TELEFAX: (510) 655-3542
; TELEX: n/a
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-529-169A-2

Query Match 6.8%; Score 8; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
| | | | |
Db 456 CMSADLEV 463

RESULT 400
US-09-483-799-2
; Sequence 2, Application US/09483799
; Patent No. 6472180
; GENERAL INFORMATION:
; APPLICANT: HOUGHTON, MICHAEL
; APPLICANT: CHOO, QUI-LIM
; APPLICANT: HAN, JANG
; APPLICANT: CHOE, JOONHO
; TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
; TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: Intellectual Property - R440, P.O. Box 8097
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: 08/529,169
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0100.005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3274
; TELEFAX: (510) 655-3542
; TELEX: n/a
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 465 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-483-799-2

Query Match 6.8%; Score 8; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
| | | | |
Db 456 CMSADLEV 463

RESULT 401
US-08-867-611-58
; Sequence 58, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A

; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; TITLE OF INVENTION: ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/614,069
; FILING DATE: 07-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,561
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,565
; FILING DATE: 21-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,566
; FILING DATE: 21-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 4834.US.P6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-937-9556
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 512 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-867-611-58

Query Match 6.8%; Score 8; DB 2; Length 512;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECSQ 67
| | | | |
Db 277 DEMEECSQ 284

RESULT 402
US-09-690-359-58
; Sequence 58, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G

DESAL, SURESH M
CASEY, JAMES M
DAILEY, STEPHEN H
DAWSON, GEORGE J
GUTIERREZ, ROBIN A
LESNIEWSKI, RICHARD R
STEWART, JAMES L
RUPPRECHT, KEVIN R
TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT ANTIGENS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: ABBOTT LABORATORIES
STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
CITY: ABBOTT PARK
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/690,359
FILING DATE: 17-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/867,611
FILING DATE: 02-JUN-1997
APPLICATION NUMBER: US/08/646,757
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/179,896
FILING DATE: <Unknown>
APPLICATION NUMBER: US 07/572,822
FILING DATE: 24-AUG-1990
APPLICATION NUMBER: US 07/614,069
FILING DATE: 07-NOV-1990
APPLICATION NUMBER: US 07/748,561
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,565
FILING DATE: 21-AUG-1991
APPLICATION NUMBER: US 07/748,566
FILING DATE: 21-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: POREMSKI, PRISCILLA E
REGISTRATION NUMBER: 33,207
REFERENCE/DOCKET NUMBER: 4834.US.P6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 708-937-6365
TELEFAX: 708-937-9556
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 512 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 58:

Query Match 6.8%; Score 8; DB 2; Length 512;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMERCSQ 67
DB 277 DEMERCSQ 284

RESULT 403
US-10-104-966-6
; Sequence 6, Application US/10104966

Patent No. 6580059
GENERAL INFORMATION:
APPLICANT: Matti Sallberg
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
FILE REFERENCE: TRIPEP.23AUSC1
CURRENT APPLICATION NUMBER: US/10/104,966
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 613
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis C virus NS3 protein sequence
US-10-104-966-6

Query Match 6.8%; Score 8; DB 2; Length 613;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CWSADLEV 9
DB 604 CWSADLEV 611

RESULT 404
US-09-929-955-6
; Sequence 6, Application US/09929955
; Patent No. 6858590
GENERAL INFORMATION:
APPLICANT: Matti Sallberg
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
FILE REFERENCE: TRIPEP.23AUS2
CURRENT APPLICATION NUMBER: US/09/929,955
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 613
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis C virus NS3 protein sequence
US-09-929-955-6

Query Match 6.8%; Score 8; DB 2; Length 613;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CWSADLEV 9
DB 604 CWSADLEV 611

RESULT 405
US-08-833-678A-1
; Sequence 1, Application US/08833678A
; Patent No. 5989905

;; GENERAL INFORMATION:
;; APPLICANT: HOUGHTON, MICHAEL
;; APPLICANT: CHOO, QUI-LIM
;; APPLICANT: HAN, JANG
;; APPLICANT: CHOE, JOONHO
;; TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
;; TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Intellectual Property - R440, P.O. Box 8097
;; CITY: Emeryville
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 94662-8097
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: US/08/833,678A
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/529,169
;; FILING DATE: 15-SEP-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Harbin, Alisa A.
;; REGISTRATION NUMBER: 33,895
;; REFERENCE/DOCKET NUMBER: 0100.005
;; TELEPHONE: (510) 923-3274
;; TELEFAX: (510) 655-3542
;; TELEX: n/a
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 631 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-833-678A-1

Query Match 6.8%; Score 8; DB 1; Length 631;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 406
US-09-128-314-2
; Sequence 2, Application US/09128314
; Patent No. 6183121
;; GENERAL INFORMATION:
;; APPLICANT: Kim, Joseph L
;; APPLICANT: Morgenstern, Kurt A
;; APPLICANT: Caron, Paul R
;; APPLICANT: Lin, Chao
;; TITLE OF INVENTION: Crystall Structure Of The HCV NS3 Helicase Domain
;; FILE REFERENCE: Sequence listing for VPI/97-101
;; Patent No. 6183121
;; CURRENT APPLICATION NUMBER: US/09/128,314
;; EARLIER FILING DATE: 1998-08-03
;; EARLIER APPLICATION NUMBER: 60/055,772
;; EARLIER FILING DATE: 1997-08-13
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 2

;; LENGTH: 631
;; TYPE: PRT
;; ORGANISM: Hepatitis C virus
US-09-128-314-2

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 407
US-08-529-169A-1
; Sequence 1, Application US/08529169A
; Patent No. 6194140
;; GENERAL INFORMATION:
;; APPLICANT: HOUGHTON, MICHAEL
;; APPLICANT: CHOO, QUI-LIM
;; APPLICANT: HAN, JANG
;; APPLICANT: CHOE, JOONHO
;; TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
;; TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Chiron Corporation
;; STREET: Intellectual Property - R440, P.O. Box 8097
;; CITY: Emeryville
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 94662-8097
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 15-SEP-1995
;; CLASSIFICATION: 4325
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Harbin, Alisa A.
;; REGISTRATION NUMBER: 33,895
;; REFERENCE/DOCKET NUMBER: 0100.005
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 923-3274
;; TELEFAX: (510) 655-3542
;; TELEX: n/a
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 631 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-529-169A-1

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 408
US-09-483-799-1
; Sequence 1, Application US/09483799
; Patent No. 6472180
;; GENERAL INFORMATION:

APPLICANT: HOUGHTON, MICHAEL
APPLICANT: CHOO, QUI-LIM
APPLICANT: HAN, JANG
APPLICANT: CHOE, JOONHO
TITLE OF INVENTION: HCV NS3 PROTEIN FRAGMENTS HAVING
TITLE OF INVENTION: HELICASE ACTIVITY AND IMPROVED SOLUBILITY
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESS: CHIRON CORPORATION
STREET: Intellectual Property - R440, P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/483,799
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/529,169
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0100.005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-3274
TELEFAX: (510) 655-3542
TELEX: n/a
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-483-799-1

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 409
US-09-929-955-29
Sequence 29, Application US/09929955
Patent No. 6858590
GENERAL INFORMATION:
APPLICANT: Matti Salberg
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: TRIPEP.23AUS2
CURRENT APPLICATION NUMBER: US/09/929,955
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 29

LENGTH: 632
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis C virus NS3 peptide
US-09-929-955-29

Query Match 6.8%; Score 8; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db 623 CMSADLEV 630

RESULT 410
US-09-930-591-12
Sequence 12, Application US/09930591
Patent No. 6960569
GENERAL INFORMATION:
APPLICANT: Matti Salberg
TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
TITLE OF INVENTION: NS3/4A FUSION GENE
FILE REFERENCE: TRIPEP.028AUS
CURRENT APPLICATION NUMBER: US/09/930,591
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 632
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis C virus NS3 peptide
US-09-930-591-12

Query Match 6.8%; Score 8; DB 2; Length 632;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db 623 CMSADLEV 630

RESULT 411
US-09-288-391-25
Sequence 25, Application US/09288391
Patent No. 6251583
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Malcolm, Bruce
APPLICANT: Beyer, Brian
APPLICANT: Njoroge, George
APPLICANT: Durkin, James
APPLICANT: Windsor, William
TITLE OF INVENTION: No. 6251583el Peptide Substrates for HCV NS3 Protease Assay
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/288,391
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McLaughlin, Jaye P.
;; REGISTRATION NUMBER: 41,211
;; REFERENCE/DOCKET NUMBER: IN0829P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (908)298-5056
;; TELEFAX: (908)298-5388
;; INFORMATION FOR SEQ ID NO: 25:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 638 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-288-391-25

Query Match 6.8%; Score 8; DB 2; Length 638;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 629 CMSADLEV 636
|||||||

RESULT 412
US-08-867-611-4
; Sequence 4, Application US/08867611
; Patent No. 6172189
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESS: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,611
; FILING DATE: 02-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,757
; FILING DATE:
; APPLICATION NUMBER: US/08/179,896
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/572,822

;; FILING DATE: 24-AUG-1990
;; PRIOR APPLICATION DATA: US 07/614,069
;; APPLICATION NUMBER: 07-NOV-1990
;; FILING DATE: 07-NOV-1990
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,561
;; FILING DATE: 21-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,565
;; FILING DATE: 21-AUG-1991
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/748,566
;; FILING DATE: 21-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834.US.P6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 781 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-867-611-4
Query Match 6.8%; Score 8; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMBEC5Q 67
Db 546 DEMBEC5Q 553
|||||||
RESULT 413
US-09-690-359-4
; Sequence 4, Application US/09690359
; Patent No. 6593083
; GENERAL INFORMATION:
; APPLICANT: DEVARE, SUSHIL G
; APPLICANT: DESAI, SURESH M
; APPLICANT: CASEY, JAMES M
; APPLICANT: DAILEY, STEPHEN H
; APPLICANT: DAWSON, GEORGE J
; APPLICANT: GUTIERREZ, ROBIN A
; APPLICANT: LESNIEWSKI, RICHARD R
; APPLICANT: STEWART, JAMES L
; APPLICANT: RUPPRECHT, KEVIN R
; TITLE OF INVENTION: HEPATITIS C ASSAY UTILIZING RECOMBINANT
; ANTIGENS
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESS: ABBOTT LABORATORIES
; STREET: ONE ABBOTT PARK ROAD, CHAD377/AP6D2
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/690,359
; FILING DATE: 17-Oct-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/867,611
; FILING DATE: 02-JUN-1997

;; APPLICATION NUMBER: US/08/646,757
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: US/08/179,896
;; FILING DATE: <Unknown>
;; APPLICATION NUMBER: US 07/572,822
;; FILING DATE: 24-AUG-1990
;; APPLICATION NUMBER: US 07/614,069
;; FILING DATE: 07-NOV-1990
;; APPLICATION NUMBER: US 07/748,561
;; FILING DATE: 21-AUG-1991
;; APPLICATION NUMBER: US 07/748,565
;; FILING DATE: 21-AUG-1991
;; APPLICATION NUMBER: US 07/748,566
;; FILING DATE: 21-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834.US.P6
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 781 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-690-359-4

Query Match 6.8%; Score 8; DB 2; Length 781;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
Db 546 DEMECSQ 553

RESULT 414
PCT-US92-06965A-9
;; Sequence 9, Application PC/TUS9206965A
;; GENERAL INFORMATION:
;; APPLICANT: DEVARE, S.
;; APPLICANT: DESAI, S.
;; APPLICANT: DAILEY, S.
;; TITLE OF INVENTION: HCV SYNTHETIC PEPTIDE FROM NS1 REGION
;; NUMBER OF SEQUENCES: 35
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ABBOTT LABORATORIES
;; STREET: ONE ABBOTT PARK ROAD
;; CITY: ABBOTT PARK
;; STATE: ILLINOIS
;; COUNTRY: U.S.
;; ZIP: 60065-3500
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US92/06965A
;; FILING DATE: 19920821
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: FOREMSKI, PRISCILLA E.
;; REGISTRATION NUMBER: 33,207
;; REFERENCE/DOCKET NUMBER: 4834PC.02
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 708-937-6365
;; TELEFAX: 708-937-9556
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 781 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
PCT-US92-06965A-9

Query Match 6.8%; Score 8; DB 4; Length 781;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMECSQ 67
Db 546 DEMECSQ 553

RESULT 415
US-08-444-818-54
;; Sequence 54, Application US/08444818
;; Patent No. 6150087
;; GENERAL INFORMATION:
;; APPLICANT: Chien, David Y.
;; APPLICANT: Rutter, William J.
;; TITLE OF INVENTION: NANV Diagnostics and Vaccines
;; NUMBER OF SEQUENCES: 777
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Chiron Corporation
;; STREET: 4560 Horton Street
;; CITY: Emeryville
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94608-2916
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/444,818
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/403,590
;; FILING DATE: 14-MAR-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Harbin, Alisa A.
;; REGISTRATION NUMBER: 33,895
;; REFERENCE/DOCKET NUMBER: 0110.002
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (508)359-3876
;; TELEFAX: (508)359-3885
;; INFORMATION FOR SEQ ID NO: 54:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1786 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-444-818-54

Query Match 6.8%; Score 8; DB 2; Length 1786;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db 932 CMSADLEV 939

RESULT 416
US-08-146-028-284
;; Sequence 284, Application US/08146028
;; Patent No. 5891640
;; GENERAL INFORMATION:
;; APPLICANT:

;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA: US/08/146,028
;; APPLICATION FOR SEQ ID NO: 284;
;; INFORMATION FOR SEQ ID NO: 284:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-146-028-284

Query Match 5.9%; Score 7; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECS 66
| | | | | | | | |
Db 3 DEMEECS 9

RESULT 417

US-08-146-028-290
; Sequence 290, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

APPLICANT:

;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; INFORMATION FOR SEQ ID NO: 290:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-146-028-290

Query Match 5.9%; Score 7; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECS 66
| | | | | | | | |
Db 3 DEMEECS 9

RESULT 418

US-08-146-028-293
; Sequence 293, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA: US/08/146,028
;; APPLICATION FOR SEQ ID NO: 293;
;; INFORMATION FOR SEQ ID NO: 293:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-146-028-293

Query Match 5.9%; Score 7; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMEBCSQ 67
| | | | | | | | |
Db 1 EMEBCSQ 7

RESULT 419

US-08-146-028-299
; Sequence 299, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:

APPLICANT:

;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; INFORMATION FOR SEQ ID NO: 299:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-146-028-299

Query Match 5.9%; Score 7; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMEBCSQ 67
| | | | | | | | |
Db 1 EMEBCSQ 7

RESULT 420

US-08-802-981-212
; Sequence 212, Application US/08802981
; Patent No. 6037137

```

; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-0003000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 212:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-802-981-212

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
DB 1 EMECSQ 7

RESULT 421
US-08-723-425A-284
; Sequence 284, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 290:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-723-425A-290

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 284:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-723-425A-284

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMBECs 66
DB 3 DEMBECs 9

RESULT 422
US-08-723-425A-290
; Sequence 290, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 290:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-723-425A-290

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 60 DEMECS 66
| | | | |
Db 3 DEMECS 9

RESULT 423
US-08-723-425A-293
; Sequence 293, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723.425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 293:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-723-425A-293

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMEECSQ 67
| | | | |
Db 1 EMEECSQ 7

RESULT 424
US-08-723-425A-299
; Sequence 299, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHUYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA

COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723.425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 299:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-723-425A-299

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMEECSQ 67
| | | | |
Db 1 EMEECSQ 7

RESULT 425
US-09-112-206-284
; Sequence 284, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112.206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 284:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-284

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMECS 66

```
Db          |||||
            3 DEMECS 9

RESULT 426
US-09-112-206-290
; Sequence 290, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 290:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-290

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          60 DEMECS 66
            |||||
            3 DEMECS 9

Db          |||||

RESULT 427
US-09-112-206-293
; Sequence 293, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 293:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; MOLECULE TYPE: peptide
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-293

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          61 EMECSQ 67
            |||||
            1 EMECSQ 7

Db          |||||

RESULT 428
US-09-112-206-299
; Sequence 299, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 299:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-299

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          61 EMECSQ 67
            |||||
            1 EMECSQ 7

Db          |||||

RESULT 429
US-09-790-497A-274
; Sequence 274, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
```

```
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 274
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-274

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEECS 66
Db      3 DEMEECS 9
|||||

RESULT 430
US-09-790-497A-277
; Sequence 277, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 277
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-277

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      61 EMERCSQ 67
Db      1 EMERCSQ 7
|||||

RESULT 431
US-09-576-824A-274
; Sequence 274, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
```

```
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 274
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-274

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEECS 66
Db      3 DEMEECS 9
|||||

RESULT 432
US-09-576-824A-277
; Sequence 277, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 277
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-277

Query Match          5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      61 EMERCSQ 67
Db      1 EMERCSQ 7
|||||

RESULT 433
```

US-09-680-497-284
; Sequence 284, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 284:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-284

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEECS 66
Db 3 DEMEECS 9

RESULT 434
US-09-680-497-290
; Sequence 290, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 290:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-290

Query Match 5.9%; Score 7; DB 2; Length 9;

Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEECS 66
Db 3 DEMEECS 9
RESULT 435
US-09-680-497-293
; Sequence 293, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 293:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-293

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 DEMEECSQ 67
Db 1 EMEECSQ 7

RESULT 436
US-09-680-497-299
; Sequence 299, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 299:

;; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-299

Query Match 5.9%; Score 7; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMECSQ 67
| | | | |
Db 1 EMECSQ 7

RESULT 437
US-08-617-929-20
; Sequence 20, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-617-929-20

Query Match 5.9%; Score 7; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
| | | | |

Db 4 PDKEVLY 10

RESULT 438
US-08-617-929-22
; Sequence 22, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-617-929-22

Query Match 5.9%; Score 7; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
| | | | |
Db 5 PDKEVLY 11

RESULT 439
US-08-439-747A-14
; Sequence 14, Application US/08439747A
; Patent No. 5767233
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Murray, Michael
; APPLICANT: Ramanathan, Lata
; TITLE OF INVENTION: Soluble, Cleavable Substrates of the Hepatitis
; C Protease
; NUMBER OF SEQUENCES: 34

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering Corp.
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07033-0530
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Macintosh 7.5.3
;; SOFTWARE: Microsoft Word 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/439,747A
;; FILING DATE: May 12, 1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lunn, Paul G.
;; REGISTRATION NUMBER: 32,743
;; REFERENCE/DOCKET NUMBER: JB0509
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908-298-5061
;; TELEFAX: 908-298-5388
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 14 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: polypeptide
;; FEATURE:
;; NAME/KEY: NS4A Active Mutant
;; US-08-439-747A-14

Query Match 5.9%; Score 7; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 32 GCWIVG 38
Db 1 GCWIVG 7

RESULT 440
US-08-440-409B-14
; Sequence 14, Application US/08440409B
; Patent No. 5843752
; GENERAL INFORMATION:
; APPLICANT: Dasmahapatra, Bimal
; APPLICANT: Murray, Michael
; APPLICANT: Ramanathan, Lata
; APPLICANT: Butkiewicz, Nancy
; TITLE OF INVENTION: Soluble Active Hepatitis C Virus Protease
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5.3
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,409B
; FILING DATE: May 12, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: 32,743

;; REFERENCE/DOCKET NUMBER: JB0494
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 908-298-5061
;; TELEFAX: 908-298-5388
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 14 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: polypeptide
;; FEATURE:
;; NAME/KEY: NS4A Active Mutant
;; US-08-440-409B-14

Query Match 5.9%; Score 7; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0;

QY 32 GCWIVG 38
Db 1 GCWIVG 7

RESULT 441
US-08-617-929-21
; Sequence 21, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KOMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-617-929-21

Query Match 5.9%; Score 7; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
|||||
Db 6 PDKEVLY 12

RESULT 442

US-08-244-116B-4
; Sequence 4, Application US/08244116B
; Patent No. 5763159
; GENERAL INFORMATION:
; APPLICANT: Simmonds, Peter
; APPLICANT: Chan, Shiu-Wan
; APPLICANT: Yip, Peng L.
; TITLE OF INVENTION: Hepatitis-C Virus Testing
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson, P.A.
; STREET: 1211 East Morehead Street
; CITY: Charlotte
; STATE: No. 5763159th Carolina
; COUNTRY: United States
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0. Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/244.116B
; FILING DATE: 15-JUL-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/02143
; FILING DATE: 20-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31.665
; REFERENCE/DOCKET NUMBER: 1749-125
; TELEPHONE: 704-334-2014
; TELEFAX: 704-334-1561
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis-C virus
US-08-244-116B-4

Query Match 5.9%; Score 7; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
|||||
Db 6 PDKEVLY 12

RESULT 443

US-08-802-981-96
; Sequence 96, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:

; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0. Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-0003000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "Aib"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /product= "Acp"
US-08-802-981-96

Query Match 5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMEECSQ 67
|||||
Db 6 EMEECSQ 12

RESULT 444

US-08-537-802-28
; Sequence 28, Application US/08537802
; Patent No. 6881821
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS-C VIRUS TYPE 4, 5 & 6
; NUMBER OF SEQUENCES: 50
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0. Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,802
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/00957
; FILING DATE:

APPLICATION NUMBER: GB 9309237.7
FILING DATE: 05-MAY-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 9400263.1
FILING DATE: 07-JAN-1994
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-537-802-28

Query Match 5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 6 PDKEVLY 12

RESULT 445

US-09-747-287A-139
Sequence 139, Application US/09747287A
Patent No. 6893868
GENERAL INFORMATION:
APPLICANT: KOMORIYA, AKIRA
TITLE OF INVENTION: HOMO-DOUBLY LABELED COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL SAMPLES
FILE REFERENCE: 300-948600US
CURRENT APPLICATION NUMBER: US/09/747,287A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/349,019
PRIOR FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: US08/802,981
PRIOR FILING DATE: 1997-02-20
PRIOR APPLICATION NUMBER: PCT/US00/24882
PRIOR FILING DATE: 2000-09-11
NUMBER OF SEQ ID NOS: 246
SOFTWARE: PatentIn version 3.3
SEQ ID NO 139
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic peptide.
NAME/KEY: misc_feature
LOCATION: (4)..(4)
OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-747-287A-139

Query Match 5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
DB 6 EMECSQ 12

RESULT 446

US-09-394-019C-92
Sequence 92, Application US/09394019C
Patent No. 6936687
GENERAL INFORMATION:
APPLICANT: Oncoimmunin, Inc.
APPLICANT: Komoriya, Akira
TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL SAMPLES AND METHODS OF USE THEREOF

FILE REFERENCE: 300-903820US
CURRENT APPLICATION NUMBER: US/09/394,019C
CURRENT FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: PCT/US98/00300
PRIOR FILING DATE: 1998-02-20
PRIOR APPLICATION NUMBER: US 08/802,981
PRIOR FILING DATE: 1997-02-20
NUMBER OF SEQ ID NOS: 405
SOFTWARE: PatentIn version 3.2
SEQ ID NO 92
LENGTH: 18
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide substrate
NAME/KEY: MOD_RES
LOCATION: (3)..(3)
OTHER INFORMATION: X is Alb
NAME/KEY: misc_feature
LOCATION: (3)..(4)
OTHER INFORMATION: Xaa can be any naturally occurring amino acid
NAME/KEY: MOD_RES
LOCATION: (4)..(4)
OTHER INFORMATION: X is epsilon-aminocaproic acid
US-09-394-019C-92

Query Match 5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
DB 6 EMECSQ 12

RESULT 447

US-09-394-019C-337
Sequence 337, Application US/09394019C
Patent No. 6936687
GENERAL INFORMATION:
APPLICANT: Oncoimmunin, Inc.
APPLICANT: Komoriya, Akira
TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL SAMPLES AND METHODS OF USE THEREOF
FILE REFERENCE: 300-903820US
CURRENT APPLICATION NUMBER: US/09/394,019C
CURRENT FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: PCT/US98/00300
PRIOR FILING DATE: 1998-02-20
PRIOR APPLICATION NUMBER: US 08/802,981
PRIOR FILING DATE: 1997-02-20
NUMBER OF SEQ ID NOS: 405
SOFTWARE: PatentIn version 3.2
SEQ ID NO 337
LENGTH: 18
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic peptide. Chemically synthesized protease substrate.
NAME/KEY: misc_feature
LOCATION: (3)..(3)
OTHER INFORMATION: Xaa is alpha-aminoisobutyric acid
NAME/KEY: misc_feature
LOCATION: (4)..(4)
OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-394-019C-337

```
Query Match          5.9%; Score 7; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMECSQ 67
Db 6 EMECSQ 12

RESULT 448
US-08-802-981-99
; Sequence 99, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-00030005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "Aib"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /product= "Acp"
US-08-802-981-99

Query Match          5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMECSQ 67
Db 6 EMECSQ 12

RESULT 449
US-08-604-365-11
; Sequence 11, Application US/08604365
; Patent No. 6183949
; GENERAL INFORMATION:
; APPLICANT: Seidel, Christoph; Ehrlich-Weinreich,
; Gertraud; Bayer, Hubert; Wienhues, Ursula; Jung,
; Gunther-Gerhard; Ihlenfeldt, Hans Georg
; TITLE OF INVENTION: HCV Peptide Antigens and Methods for
; the Determination of HCV
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
```

```
; APPLICANT: Seidel, Christoph; Ehrlich-Weinreich,
; APPLICANT: Gertraud; Bayer, Hubert; Wienhues, Ursula; Jung,
; APPLICANT: Gunther-Gerhard; Ihlenfeldt, Hans Georg
; TITLE OF INVENTION: HCV Peptide Antigens and Methods for
; the Determination of HCV
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/604,365
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/977,398
; FILING DATE: 11-MARCH-1993
; APPLICATION NUMBER: PCT/EP92/01468
; FILING DATE: 30-JUNE-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 41 22 160.5
; FILING DATE: 04-JULY-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 41 41 304.0
; FILING DATE: 14-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 09 216.9
; FILING DATE: 21-MARCH-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: BOER 1010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-604-365-11

Query Match          5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
Db 4 TNWQKLE 10

RESULT 450
US-09-689-678-11
; Sequence 11, Application US/09689678
; Patent No. 6592871
; GENERAL INFORMATION:
; APPLICANT: Seidel, Christoph; Ehrlich-Weinreich,
; Gertraud; Bayer, Hubert; Wienhues, Ursula; Jung,
; Gunther-Gerhard; Ihlenfeldt, Hans Georg
; TITLE OF INVENTION: HCV Peptide Antigens and Methods for
; the Determination of HCV
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
```

STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/689,678
FILING DATE: 13-Oct-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/604,365
FILING DATE: <Unknown>
APPLICATION NUMBER: 07/977,398
FILING DATE: 11-MARCH-1993
APPLICATION NUMBER: PCT/EP92/01468
FILING DATE: 30-JUNE-1992
APPLICATION NUMBER: DE 41 22 160.5
FILING DATE: 04-JULY-1991
APPLICATION NUMBER: DE 41 41 304.0
FILING DATE: 14-DEC-1991
APPLICATION NUMBER: DE 42 09 216.9
FILING DATE: 21-MARCH-1992
ATTORNEY/AGENT INFORMATION:
NAME: Teai, Christine H.
REGISTRATION NUMBER: 34,266
REFERENCE/DOCKET NUMBER: BOER 1010
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 19 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-689-678-11

Query Match 5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNNQKLE 112
DB 4 TNNQKLE 10

RESULT 451
US-09-747-287A-142
; Sequence 142, Application US/09747287A
; Patent No. 6893868
; GENERAL INFORMATION:
; APPLICANT: KOMORIYA, AKIRA
; APPLICANT: PACKARD, BEVERLY S.
; TITLE OF INVENTION: HOMO-DOUBLY LABELED COMPOSITIONS FOR THE DETECTION OF ENZYME
; FILE REFERENCE: 300-948600US
; CURRENT APPLICATION NUMBER: US/09/747,287A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/349,019
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US08/802,981
; PRIOR FILING DATE: 1997-02-20
; PRIOR APPLICATION NUMBER: PCT/US00/24882
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 142

LENGTH: 19
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic peptide.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4)..(4)
OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-747-287A-142

Query Match 5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
DB 6 EMECSQ 12

RESULT 452
US-09-394-019C-95
; Sequence 95, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903820US
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 95
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide substrate
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (3)..(3)
; OTHER INFORMATION: X is Aib
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(4)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)..(4)
; OTHER INFORMATION: X is epsilon aminocaproic acid
US-09-394-019C-95

Query Match 5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
DB 6 EMECSQ 12

RESULT 453
US-09-394-019C-340
; Sequence 340, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:

APPLICANT: Oncoimmunin, Inc.
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly
TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
FILE REFERENCE: 300-903820US
CURRENT APPLICATION NUMBER: US/09/394,019C
CURRENT FILING DATE: 1999-09-10
PRIOR APPLICATION NUMBER: PCT/US98/00300
PRIOR FILING DATE: 1998-02-20
PRIOR APPLICATION NUMBER: US 08/802,981
PRIOR FILING DATE: 1997-02-20
NUMBER OF SEQ ID NOS: 405
SOFTWARE: PatentIn version 3.2
SEQ ID NO 340
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: Synthetic peptide. Chemically synthesized protease substrate.
FEATURE:
NAME/KEY: misc feature
LOCATION: (3)-(3)
OTHER INFORMATION: Xaa is alpha-aminoisobutyric acid
FEATURE:
NAME/KEY: misc feature
LOCATION: (4)-(4)
OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-394-019C-340

Query Match 5.9%; Score 7; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMECSQ 67
Db 6 EMECSQ 12

RESULT 454
US-08-617-929-2
Sequence 2, Application US/08617929
Patent No. 5885771
GENERAL INFORMATION:
APPLICANT: KUMAZAWA, Toshiaki
TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
TITLE OF INVENTION: IMMUNOASSAY
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,929
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/01823
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6/207695
FILING DATE: 31-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5/272864
FILING DATE: 29-OCT-1993

ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 77384/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-929-2

Query Match 5.9%; Score 7; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 9 PDKEVLY 15

RESULT 455
US-08-802-981-102
Sequence 102, Application US/08802981
Patent No. 6037137
GENERAL INFORMATION:
APPLICANT: Komoriya, Akira
APPLICANT: Packard, Beverly S.
TITLE OF INVENTION: Compositions for the Detection of Enzyme
TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/802,981
FILING DATE: 20-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 016865-0003000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /product= "Aib"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /product= "Acp"

```
;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 16
; OTHER INFORMATION: /product= "Acp"
; US-08-802-981-102

Query Match          5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
    |||||
Db 6 EMECSQ 12

RESULT 456
US-09-790-497A-92
; Sequence 92, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-92

Query Match          5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
    |||||
Db 9 PDKEVLY 15

RESULT 457
US-09-790-497A-93
; Sequence 93, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
```

```
;
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-93
```

```
Query Match          5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 50 PDKEVLY 56
    |||||
Db 3 PDKEVLY 9
```

RESULT 458

```
US-09-790-497A-149
; Sequence 149, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-149
```

```
Query Match          5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 50 PDKEVLY 56
    |||||
Db 9 PDKEVLY 15
```

RESULT 459

```
US-09-576-824A-149
; Sequence 149, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
```

; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 149
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-149

Query Match 5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 9 PDKEVLY 15

RESULT 460

US-09-878-281A-278
; Sequence 278, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, proph
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35
; CURRENT APPLICATION NUMBER: US/09/878,281A
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-09-878-281A-278

Query Match 5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 9 PDKEVLY 15

RESULT 461

US-09-747-287A-145
; Sequence 145, Application US/09747287A
; Patent No. 6893868
; GENERAL INFORMATION:
; APPLICANT: KOMORIYA, AKIRA
; APPLICANT: PACKARD, BEVERLY S.
; TITLE OF INVENTION: HOMO-DOUBLY LABELED COMPOSITIONS FOR THE DETECTION OF ENZYME
; TITLE OF INVENTION: ACTIVITY IN BIOLOGICAL SAMPLES
; FILE REFERENCE: 300-948600US

; CURRENT APPLICATION NUMBER: US/09/747,287A
; CURRENT FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/349,019
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US08/802,981
; PRIOR FILING DATE: 1997-02-20
; PRIOR APPLICATION NUMBER: PCT/US00/24882
; PRIOR FILING DATE: 2000-09-11
; NUMBER OF SEQ ID NOS: 246
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 145
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-747-287A-145

Query Match 5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 BMECSQ 67
Db 6 BMECSQ 12

RESULT 462

US-09-394-019C-98
; Sequence 98, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; TITLE OF INVENTION: SAMPLES AND METHODS OF USE THEREOF
; FILE REFERENCE: 300-903820US
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 98
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide substrate
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (3)..(3)
; OTHER INFORMATION: X is Aib
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)..(4)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)..(4)
; OTHER INFORMATION: X is epsilon aminocaproic acid
; FEATURE:

; NAME/KEY: MOD.RES
; LOCATION: (16)..(16)
; OTHER INFORMATION: X is epsilon-aminocaproic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-09-394-019C-98

Query Match 5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
|||||
DB 6 EMECSQ 12

RESULT 463

US-09-394-019C-343
; Sequence 343, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncoimmun, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903820US
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 343
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide. Chemically synthesized protease substrate.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is alpha-aminoisobutyric acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (16)..(16)
; OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-394-019C-343

Query Match 5.9%; Score 7; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMECSQ 67
|||||
DB 6 EMECSQ 12

RESULT 464

US-08-146-028-92
; Sequence 92, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES

; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
US-08-146-028-92

Query Match 5.9%; Score 7; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
DB 10 PDKEVLY 16

RESULT 465

US-08-146-028-93
; Sequence 93, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
US-08-146-028-93

```
Query Match      5.9%; Score 7; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 4 PDKEVLY 10

RESULT 466
US-08-146-028-149
; Sequence 149, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 43
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 1
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 22
US-08-146-028-149

Query Match      5.9%; Score 7; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 10 PDKEVLY 16

RESULT 467
US-08-845-926-11
; Sequence 11, Application US/08845926
; Patent No. 5935778
; GENERAL INFORMATION:
; APPLICANT: SEIDEL, Christoph
; APPLICANT: WEINHUES-THELEN, Ursula-Henrike
; APPLICANT: SCHMITT, Urban
; APPLICANT: JUNG, G nther-Gerhard
; APPLICANT: IHLENFELDT, HANS-Georg
; APPLICANT: KRAAS, Wolfgang
; TITLE OF INVENTION: Method for serological typing using
; TITLE OF INVENTION: type-specific antigens
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
```

```
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,926
; FILING DATE: 04-28-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/598,993
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 04 302.2
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C Virus
US-08-845-926-11

Query Match      5.9%; Score 7; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
Db 15 TNWQKLE 21

RESULT 468
US-08-845-926-12
; Sequence 12, Application US/08845926
; Patent No. 5935778
; GENERAL INFORMATION:
; APPLICANT: SEIDEL, Christoph
; APPLICANT: WEINHUES-THELEN, Ursula-Henrike
; APPLICANT: SCHMITT, Urban
; APPLICANT: JUNG, G nther-Gerhard
; APPLICANT: IHLENFELDT, HANS-Georg
; APPLICANT: KRAAS, Wolfgang
; TITLE OF INVENTION: Method for serological typing using
; TITLE OF INVENTION: type-specific antigens
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/845,926
; FILING DATE: 04-28-1997
```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/598,993
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 195 04 302.2
FILING DATE: 09-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Murray, Robert B.
REGISTRATION NUMBER: 22,980
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Hepatitis C Virus
US-08-845-926-12

Query Match 5.9%; Score 7; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 TNNQKLE 112
DB 15 TNNQKLE 21

RESULT 469

US-08-723-425A-92
Sequence 92, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV type 2
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-723-425A-92

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
DB 10 PDKEVLY 16

RESULT 470

US-08-723-425A-93
Sequence 93, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:
APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHVE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV type 2
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-723-425A-93

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
   |||||
Db 4 PDKEVLY 10

RESULT 471
US-08-723-425A-149
; Sequence 149, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 1
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 22
US-08-723-425A-149

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
   |||||
Db 10 PDKEVLY 16

RESULT 472
US-09-112-206-92
; Sequence 92, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
```

```
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
US-09-112-206-92

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
   |||||
Db 10 PDKEVLY 16

RESULT 473
US-09-112-206-93
; Sequence 93, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
```

```
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 22
; US-09-112-206-93
;
; Query Match
; Best Local Similarity 100.0%; Score 7; DB 2; Length 22;
; Mismatches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 50 PDKEVLY 56
; DB 4 PDKEVLY 10
;
; RESULT 474
; US-09-112-206-149
; Sequence 149, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 1
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 22
; US-09-112-206-149
;
; Query Match
; Best Local Similarity 100.0%; Score 7; DB 2; Length 22;
; Mismatches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 50 PDKEVLY 56
; DB 10 PDKEVLY 16
;
; RESULT 475
; US-09-351-296-11
; Sequence 11, Application US/09351296
; Patent No. 6447992
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: HCV type 2
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 1
; FEATURE:
; NAME/KEY: Xaa is absent
; LOCATION: 22
; US-09-112-206-149
```

```
; APPLICANT: SEIDEL, Christoph
; APPLICANT: WEINHUES-THELEN, Ursula-Henrike
; APPLICANT: SCHMITT, Urban
; APPLICANT: JUNG, G nther-Gerhard
; APPLICANT: IHLENFELDT, HANS-Georg
; APPLICANT: KRAAS, Wolfgang
; TITLE OF INVENTION: Method for serological typing using
; TITLE OF INVENTION: type-specific antigens
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/351,296
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,926
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 04 302.2
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C Virus
; US-09-351-296-11
;
; Query Match
; Best Local Similarity 100.0%; Score 7; DB 2; Length 22;
; Mismatches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 106 TNWQKLE 112
; DB 15 TNWQKLE 21
;
; RESULT 476
; US-09-351-296-12
; Sequence 12, Application US/09351296
; Patent No. 6447992
; GENERAL INFORMATION:
; APPLICANT: SEIDEL, Christoph
; APPLICANT: WEINHUES-THELEN, Ursula-Henrike
; APPLICANT: SCHMITT, Urban
; APPLICANT: JUNG, G nther-Gerhard
; APPLICANT: IHLENFELDT, HANS-Georg
; APPLICANT: KRAAS, Wolfgang
; TITLE OF INVENTION: Method for serological typing using
; TITLE OF INVENTION: type-specific antigens
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Nikaido, Marmelstein, Murray & Oram LLP
```

```
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/351,296
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/845,926
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 04 302.2
; FILING DATE: 09-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Murray, Robert B.
; REGISTRATION NUMBER: 22,980
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; ORIGINAL SOURCE:
; ORGANISM: Hepatitis C Virus
;
US-09-351-296-12

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
Db 15 TNWQKLE 21

RESULT 477
US-09-576-824A-92
; Sequence 92, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leye, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: modified site
; NAME/KEY: VARIANT
; LOCATION: (22)
; OTHER INFORMATION: modified site
;
US-09-576-824A-92

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 4 PDKEVLY 10

RESULT 479
US-09-680-497-92
; Sequence 92, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
```

```
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: modified site
; NAME/KEY: VARIANT
; LOCATION: (22)
; OTHER INFORMATION: modified site
;
US-09-576-824A-92

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 10 PDKEVLY 16

RESULT 478
US-09-576-824A-93
; Sequence 93, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leye, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 93
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)
; OTHER INFORMATION: modified site
; NAME/KEY: VARIANT
; LOCATION: (22)
; OTHER INFORMATION: modified site
;
US-09-576-824A-93

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 4 PDKEVLY 10

RESULT 479
US-09-680-497-92
; Sequence 92, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
```

;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED-
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 92:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: HCV type 2
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 22
;; US-09-680-497-92

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
Db 10 PDKEVLY 16

RESULT 480

US-09-680-497-93
; Sequence 93, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 93:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO

;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: HCV type 2
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 22
;; US-09-680-497-93

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
Db 4 PDKEVLY 10

RESULT 481

US-09-680-497-149
; Sequence 149, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:

;; APPLICANT:
;; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
;; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
;; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
;; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
;; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
;; NUMBER OF SEQUENCES: 453
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/680,497
;; FILING DATE: 06-OCT-2000
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/146,028
;; FILING DATE: 22-NOV-1993
;; INFORMATION FOR SEQ ID NO: 149:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 22 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ORIGINAL SOURCE:
;; INDIVIDUAL ISOLATE: HCV type 2
;; FEATURE:
;; NAME/KEY: Xaa is absent
;; LOCATION: 1
;; FEATURE:
;; NAME/KEY: Xaa is absent
;; LOCATION: 22
;; US-09-680-497-149

Query Match 5.9%; Score 7; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
Db 10 PDKEVLY 16

RESULT 482

US-09-020-846-44
; Sequence 44, Application US/09020846
; Patent No. 6322965
; GENERAL INFORMATION:

ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-336-553A-19

Query Match 5.9%; Score 7; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
DB 8 PDKEVLY 14

RESULT 485
US-08-439-157-19
Sequence 19, Application US/08439157
Patent No. 6416944
GENERAL INFORMATION:
APPLICANT: CHIEN, DAVID Y.
KUO, GEORGE
TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND REAGENTS FOR USE THEREIN
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,157
FILING DATE: 11-May-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553A
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-08-439-157-19

Query Match 5.9%; Score 7; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
DB 8 PDKEVLY 14
RESULT 486
US-09-437-895-19
Sequence 19, Application US/09437895
Patent No. 6416946
GENERAL INFORMATION:
APPLICANT: CHIEN, DAVID Y.
KUO, GEORGE
TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND REAGENTS FOR USE THEREIN
NUMBER OF SEQUENCES: 75
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FORSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/437,895
FILING DATE: 09-No. 6416946-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-437-895-19

Query Match 5.9%; Score 7; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
DB 8 PDKEVLY 14
RESULT 487
US-08-934-741A-3
Sequence 3, Application US/08934741A
Patent No. 5977298
GENERAL INFORMATION:
APPLICANT: Shibata, Kenji
Yamasaki, Motoo

APPLICANT: Hamada, Masako
APPLICANT: Tamaoki, Tatsuya
APPLICANT: Kosaka, No. 5977298uo
APPLICANT: Sato, Soichiro
TITLE OF INVENTION: NOVEL CALCITONIN DERIVATIVES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Antonelli, Terry, Stout, & Kraus, LLP
STREET: Suite 1800, 1300 No. 5977298th Seventeenth Street
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22209
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM-PC
OPERATING SYSTEM: DOS
SOFTWARE: Word Perfect; Version #5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,741A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 061026/95
FILING DATE: 20-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Solomon, William I.
REGISTRATION NUMBER: 28,565
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-312-6600
TELEFAX: 703-312-6666
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 34 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1 and 9
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: /label= Xaa at location 1 is
OTHER INFORMATION: Gly crosslinked to Xaa which is Asp at location 9
NAME/KEY: Modified-site
LOCATION: 34
IDENTIFICATION METHOD: by experiment
OTHER INFORMATION: /label= Xaa at location 34
OTHER INFORMATION: /note= L-prolinamide
US-08-934-741A-3
Query Match 5.9%; Score 7; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;
Qy 20 VLALAA 26
Db 10 VLALAA 16
RESULT 488
US-07-946-054-8
Sequence 8, Application US/07946054
Patent No. 5582968
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
APPLICANT: Hosein, Barbara H
TITLE OF INVENTION: No. 5582968el Branched Hybrid and Cluster
TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-A,
TITLE OF INVENTION: No. 5582968-B Hepatitis
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:

ADDRESSEE: United Biomedical Inc.
STREET: 25 Davids Dr.
CITY: Hauppauge
STATE: New York
COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/946,054
FILING DATE: 15-SEP-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-273-2828
TELEFAX: 516-273-1717
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-946-054-8
Query Match 5.9%; Score 7; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 13; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;
Qy 106 TNWQLE 112
Db 24 TNWQLE 30
RESULT 489
US-08-530-550-35
Sequence 35, Application US/08530550
Patent No. 5736321
GENERAL INFORMATION:
APPLICANT: Hosein, Barbara
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: Peptides Effective for Diagnosis and
TITLE OF INVENTION: Detection of Hepatitis c Infection
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: M. Lisa Wilson
STREET: 25 Davids Drive
CITY: Hauppauge
STATE: NY
COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,550
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, M. Lisa
REGISTRATION NUMBER: 34,045
REFERENCE/DOCKET NUMBER: 2000Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516)273-2828
TELEFAX: (516)273-1717
INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-530-550-35

Query Match 5.9%; Score 7; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNNQKLE 112
DB 24 TNNQKLE 30

RESULT 490

US-08-262-037-22
Sequence 22, Application US/08262037
Patent No. 5747239

GENERAL INFORMATION:

APPLICANT: Chang Yi Wang and Barbara Hosein
TITLE OF INVENTION: SYNTHETIC PEPTIDES SPECIFIC FOR
TITLE OF INVENTION: THE DETECTION OF ANTIBODIES TO HCV, DIAGNOSIS OF HCV
TITLE OF INVENTION: INFECTION AND PREVENTION THEREOF AS VACCINES
NUMBER OF SEQUENCES: 136
CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVE.
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/262,037
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/719,819
FILING DATE: 24-June-1991
APPLICATION NUMBER: 07/667,275
FILING DATE: 11-Mar-1991
APPLICATION NUMBER: 07/651,735
FILING DATE: 07-Feb-1991
APPLICATION NUMBER: 07/558,799
FILING DATE: 26-July-1990
APPLICATION NUMBER: 07/510,153
FILING DATE: 16-April-1990

ATTORNEY/AGENT INFORMATION:

NAME: Maria C. H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4043 US3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: (212) 751-6849
TELEX: 421792

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: Amino acid
STRANDEDNESS:

TOPOLOGY: Unknown

US-08-262-037-22

Query Match 5.9%; Score 7; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNNQKLE 112
DB 24 TNNQKLE 30

RESULT 491

US-08-617-929-10
Sequence 10, Application US/08617929
Patent No. 5885771

GENERAL INFORMATION:

APPLICANT: KUMAZAWA, Toshiaki
TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
TITLE OF INVENTION: IMMUNOASSAY
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,929
FILING DATE: 24-APR-1996
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: WO PCT/JP94/01823
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6/207695
FILING DATE: 31-AUG-1994

APPLICATION DATA:

APPLICATION NUMBER: JP 5/272864
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:

NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 77384/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-617-929-10

Query Match 5.9%; Score 7; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 11 PDKEVLY 17

RESULT 492

PCT-US93-08638-8

Sequence 8, Application PC/TUS9308638
GENERAL INFORMATION:

APPLICANT: United Biomedical Inc.
TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
TITLE OF INVENTION: Non-B Hepatitis
NUMBER OF SEQUENCES: 12

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: UNITED BIOMEDICAL INC.
;; STREET: 25 Davids Drive
;; CITY: Hauppauge
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 11788
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US93/08638
;; FILING DATE:
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: M. Lisa Wilson
;; REGISTRATION NUMBER: 34,045
;; REFERENCE/DOCKET NUMBER: 9055
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 516-273-2828
;; TELEFAX: 516-273-1717
;; TELEX:
;;
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 40 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US93-08638-8

Query Match 5.9%; Score 7; DB 4; Length 40;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
| | | | |
Db 24 TNWQKLE 30

RESULT 493
US-09-270-767-39363
; Sequence 39363, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39363
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39363

Query Match 5.9%; Score 7; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
| | | | |
Db 39 QPKGKVL 45

RESULT 494
US-09-270-767-54580
; Sequence 54580, Application US/09270767

;; Patent No. 6703491
;; GENERAL INFORMATION:
;; APPLICANT: Homburger et al.
;; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;; FILE REFERENCE: File Reference: 7326-094
;; CURRENT APPLICATION NUMBER: US/09/270,767
;; CURRENT FILING DATE: 1999-03-17
;; NUMBER OF SEQ ID NOS: 62517
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 54580
;; LENGTH: 53
;; TYPE: PRT
;; ORGANISM: Drosophila melanogaster
;; FEATURE:
;; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54580

Query Match 5.9%; Score 7; DB 2; Length 53;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
| | | | |
Db 39 QPKGKVL 45

RESULT 495
US-08-685-764-4
; Sequence 4, Application US/08685764
; Patent No. 5800982
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, AKIRA
; APPLICANT: MAKI, NOBORU
; APPLICANT: YAGI, SHINTARO
; APPLICANT: KASHIMAKUMA, TOMIKO
; TITLE OF INVENTION: ANTIGENIC PEPTIDES FOR GROUPING
; TITLE OF INVENTION: HEPATITIS C VIRUS, KIT COMPRISING THE SAME AND
; TITLE OF INVENTION: METHODS FOR ITS GROUPING USING THE SAME
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,764
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/092,192
; FILING DATE: 15-JUL-1993
; APPLICATION NUMBER: JP 212061/92
; FILING DATE: 16-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316634/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316635/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 104754/93
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: RUCKLEY, LINDA M.
; REGISTRATION NUMBER: 31003

REFERENCE/DOCKET NUMBER: 42822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 87 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-685-764-4

Query Match 5.9%; Score 7; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
DB 23 PDKEVLY 29

RESULT 496

US-09-248-796A-17488
Sequence 17488, Application US/09248796A
Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 17488

LENGTH: 100

TYPE: PRT

ORGANISM: Candida albicans

US-09-248-796A-17488

Query Match 5.9%; Score 7; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIELGK 45
|||||
DB 46 HIELGK 52

RESULT 497

US-09-107-532A-6600
Sequence 6600, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO

CORRESPONDENCE ADDRESS:

ADDRESS: GENOME THERAPEUTICS CORPORATION

STREET: 100 Beaver Street

CITY: Waltham

STATE: Massachusetts

COUNTRY: USA

ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-ROM ISO9660

COMPUTER: PC

OPERATING SYSTEM: <Unknown>

SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598

FILING DATE: 14 May 1998

APPLICATION NUMBER: 60/051571

FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Ariniello, Pamela Deneke

REGISTRATION NUMBER: 40,489

REFERENCE/DOCKET NUMBER: GTC-012

TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007

TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6600:

SEQUENCE CHARACTERISTICS:

LENGTH: 114 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: YES

ORIGINAL SOURCE:

ORGANISM: Enterococcus faecium

FEATURE:

NAME/KEY: misc feature

LOCATION: (B) LOCATION 1...114

SEQUENCE DESCRIPTION: SEQ ID NO: 6600:

US-09-107-532A-6600

Query Match 5.9%; Score 7; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIELGK 45
|||||
DB 108 HIELGK 114

RESULT 498

US-09-878-281A-60

Sequence 60, Application US/09878281A

Patent No. 6762024

GENERAL INFORMATION:

APPLICANT: Innogenetics N.V.

TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy

FILE REFERENCE: 35

CURRENT APPLICATION NUMBER: US/09/878,281A

CURRENT FILING DATE: 2001-06-12

NUMBER OF SEQ ID NOS: 284

SOFTWARE: PatentIn version 3.1

SEQ ID NO 60

LENGTH: 128

TYPE: PRT

ORGANISM: hepatitis C virus

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (13)...(13)

OTHER INFORMATION: "Xaa" is any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (26)...(27)

OTHER INFORMATION: "Xaa" is any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (51)...(51)

OTHER INFORMATION: "Xaa" is any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (78)...(78)

OTHER INFORMATION: "Xaa" is any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (78)...(78)

OTHER INFORMATION: "Xaa" is any amino acid

```
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (117)..(117)
; OTHER INFORMATION: "Xaa" is any amino acid
US-09-878-281A-60

Query Match          5.9%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEECS 66
      |||||
Db      61 DEMEECS 67

RESULT 499
US-09-252-991A-25536
; Sequence 25536, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25536
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25536

Query Match          5.9%; Score 7; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      7 LEVTTST 13
      |||||
Db      49 LEVTTST 55

RESULT 500
US-09-164-615-34
; Sequence 34, Application US/09164615B
; Patent No. 6747188
; GENERAL INFORMATION:
; APPLICANT: Hanley-Bowdoin, Linda
; APPLICANT: Settlage, Sharon
; TITLE OF INVENTION: Geminivirus Resistant Transgenic Plants
; FILE REFERENCE: 5051-433
; CURRENT APPLICATION NUMBER: US/09/164,615B
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TYCLIV mutant
; OTHER INFORMATION: C3 (mc3#47)
US-09-164-615-34

Query Match          5.9%; Score 7; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GVLAAAL 25

; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (117)..(117)
; OTHER INFORMATION: "Xaa" is any amino acid
US-09-602-787A-448
; Sequence 448, Application US/09602787A
; Patent No. 6696561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Krüger, Burkhard
; APPLICANT: Schöder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USSN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932180.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932182.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932190.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932191.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932212.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932227.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932228.7
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932229.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932927.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940765.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
```

; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 448
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-448

Query Match 5.9%; Score 7; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRT 94
|||||
DB 58 GLLQRT 64

RESULT 502

US-09-621-976-4722
; Sequence 4722, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 4722
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 20
; OTHER INFORMATION: Xaa = Ser,Thr
US-09-621-976-4722

Query Match 5.9%; Score 7; DB 2; Length 158;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWIVG 38
|||||
DB 7 GCWIVG 13

RESULT 503

US-09-252-991A-29832
; Sequence 29832, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29832
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29832

Query Match 5.9%; Score 7; DB 2; Length 163;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
|||||
DB 154 LLGGVLA 160

RESULT 504

US-09-687-363-10
; Sequence 10, Application US/09687363
; Patent No. 6683168
; GENERAL INFORMATION:
; APPLICANT: Rothstein, Thomas
; APPLICANT: Schneider, Thomas
; APPLICANT: Donohoe, Terrence
; TITLE OF INVENTION: NOVEL GENES AND METHODS THAT MODULATE APOPTOSIS
; FILE REFERENCE: 1586-50181 C
; CURRENT APPLICATION NUMBER: US/09/687,363
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/08658
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 60/124,805
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/082,503
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Mouse
US-09-687-363-10

Query Match 5.9%; Score 7; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 TTSTWVL 16
|||||
DB 96 TTSTWVL 102

RESULT 505

US-09-902-540-11456
; Sequence 11456, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16925
; SEQ ID NO 11456
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11456

Query Match 5.9%; Score 7; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKVL 87
Db 22 QFKGKVL 28
|||||

RESULT 506
US-09-902-540-10069
; Sequence 10069, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10069
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10069

Query Match 5.9%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 LLQRATQ 95
Db 95 LLQRATQ 101
|||||

RESULT 507
US-09-687-363-12
; Sequence 12, Application US/09687363
; Patent No. 6683168
; GENERAL INFORMATION:
; APPLICANT: Rothelein, Thomas
; APPLICANT: Schneider, Thomas
; APPLICANT: Donohoe, Terrence
; TITLE OF INVENTION: NOVEL GENES AND METHODS THAT MODULATE APOPTOSIS
; FILE REFERENCE: 1586-50181 C
; CURRENT APPLICATION NUMBER: US/09/687,363
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/08658
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 60/124,805
; PRIOR FILING DATE: 1999-03-15
; PRIOR APPLICATION NUMBER: 60/082,503
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Mouse
US-09-687-363-12

Query Match 5.9%; Score 7; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 TTSTWVL 16
Db 118 TTSTWVL 124
|||||

RESULT 508
US-09-583-110-3075
; Sequence 3075, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3075
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3075

Query Match 5.9%; Score 7; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 193 GGVLAAL 199
|||||

RESULT 509
US-09-602-787A-442
; Sequence 442, Application US/09602787A
; Patent No. 6695561
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Mark
; APPLICANT: Kruger, Burkhard
; APPLICANT: Schoder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-125CP
; CURRENT APPLICATION NUMBER: US/09/602,787A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: USN 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931454.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931478.0
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931563.9
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932122.1
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932124.8
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: DE 19932128.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932180.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932182.5
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932190.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932191.4
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932209.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932212.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932227.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932228.7
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932229.5
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932230.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19932927.3
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933005.0
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19933006.9
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19940764.9
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940765.7
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940766.5
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940830.0
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940831.9
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940832.7
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19940833.5
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19941378.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941379.7
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19941395.9
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19942077.7
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942078.5
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942079.3
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: DE 19942088.2
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 678
SEQ ID NO 442
LENGTH: 222
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-602-787A-442

Query Match 5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRT 94
|||||
Db 58 GLLQRT 64
|||||

RESULT 510

US-09-107-433-3655
Sequence 3655, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3655:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...227
SEQUENCE DESCRIPTION: SEQ ID NO: 3655:
US-09-107-433-3655

Query Match 5.9%; Score 7; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
|||||
Db 211 GGVLAAL 217
|||||

RESULT 511

US-09-248-796A-17487
Sequence 17487, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstein et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
CURRENT FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 17487
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-17487

Query Match 5.9%; Score 7; DB 2; Length 254;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELGGK 45
Db 52 HIELGGK 58
|||||

RESULT 512

US-10-104-966-8
; Sequence 8, Application US/10104966
; Patent No. 6680059

; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS4B protein sequence
US-10-104-966-8

Query Match 5.9%; Score 7; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
Db 41 TNWQKLE 47
|||||

RESULT 513

US-09-929-955-8
; Sequence 8, Application US/09929955
; Patent No. 6858590

; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS4B protein sequence
US-09-929-955-8

Query Match 5.9%; Score 7; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 TNWQKLE 112
Db 41 TNWQKLE 47
|||||

RESULT 514

US-09-134-001C-4206
; Sequence 4206, Application US/09134001C
; Patent No. 6380370

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4206
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4206

Query Match 5.9%; Score 7; DB 2; Length 263;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 192 LLGGVLA 198
|||||

RESULT 515

US-09-100-557-1
; Sequence 1, Application US/09100557
; Patent No. 6010848

; GENERAL INFORMATION:
; APPLICANT: DelVecchio, Alfred
; APPLICANT: Zhong, Weidong
; TITLE OF INVENTION: SCREENING METHODS USING AN
; TITLE OF INVENTION: ATPASE PROTEIN FROM A VIRUS OF THE FLAVIVIRIDAE FAMILY
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,557
; FILING DATE: 19-JUN-1998

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/051,582
FILING DATE: 02-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Dinner, Dara L.
REGISTRATION NUMBER: 33,680
REFERENCE/DOCKET NUMBER: F50675
TELEPHONE: 610-270-5017
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 269 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-100-557-1

Query Match 5.9%; Score 7; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TNNQKLE 112
DB 41 TNNQKLE 47

RESULT 516
US-08-403-852D-18
Sequence 18, Application US/08403852D
Patent No. 5891695

GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-403-852D-18

Query Match 5.9%; Score 7; DB 1; Length 277;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 240 VLAALAA 246

RESULT 517

US-08-510-646B-19
Sequence 19, Application US/08510646B
Patent No. 6077699

GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanc, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,646B
FILING DATE: 03-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,852
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 277 amino acids

;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-510-646B-19

Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 240 VLAALAA 246

RESULT 518
US-09-231-818-18
; Sequence 18, Application US/09231818
; Patent No. 6171846

;
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanc, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/231.818
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403.852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-09-231-818-18
Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 240 VLAALAA 246

RESULT 519
US-09-635-359B-18
; Sequence 18, Application US/09635359B
; Patent No. 6670157

;
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanc, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/635.359B
; FILING DATE: 09-Aug-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/231.818
; FILING DATE: 15-JAN-1999
; APPLICATION NUMBER: US 08/403.852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 277 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-09-635-359B-18
Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-635-359B-18
Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 240 VLAALAA 246

ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/020,846
FILING DATE: 09-FEB-1998
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 9-027015
FILING DATE: 10-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-024045
FILING DATE: 09-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Wegner, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 053466/0225
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 396 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-020-846-69

```
Query Match      5.9%; Score 7; DB 2; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 50 PDKEVLY 56
|||
pb 339 PDKEVLY 345

```

RESULT 524
US-10-133-007-11
; Sequence 11, Application US/10133007
; Patent No. 6623921
; GENERAL INFORMATION:
; APPLICANT: Aoyagi, Katsumi
; APPLICANT: Ohue, Chiharu
; APPLICANT: Iida, Kumiko
; APPLICANT: Yagi, Shintaro
; TITLE OF INVENTION: METHOD FOR MEASUREMENT OF HEPATITIS C VIRUS
; FILE REFERENCE: 594.352USMO
; CURRENT APPLICATION NUMBER: US/10/133,007
; CURRENT FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US/09/509,449
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: JP 10-216094
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: PCT/JP99/04129
; PRIOR FILING DATE: 1999-07-30
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 396
; TYPE: PR1
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence coding for chimeric antigen.

```

```

US-10-133-007-11

Query Match          5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 1.1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      50 PDKEVLY 56
      |||||
Db      339 PDKEVLY 345

RESULT 525
US-09-252-991A-30461
; Sequence 30461, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30461
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30461

```

```
Query Match      5.9%; Score 7; DB 2; Length 409;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 18 GGVLAAAL 24
|||
Db 160 GGVLAAAL 166

RESULT 526
US-09-020-846-68
; Sequence 68, Application US/09020846
; Patent No. 6322965
; GENERAL INFORMATION:
; APPLICANT: YAMAGUCHI, Kenjiro
; APPLICANT: KASHIWAKUMA, Tomiko
; APPLICANT: CHIBA, Yukie
; APPLICANT: YAGI, Shintaro
; APPLICANT: HASEGAWA, Akira
; TITLE OF INVENTION: CHIMERA ANTIGEN PEPTIDE
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,846
; FILING DATE: 09-FEB-1998
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 9-027015
; FILING DATE: 10-FEB-1997

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-024045
; FILING DATE: 09-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Wegner, Harold C.
; REGISTRATION NUMBER: 25,258
; REFERENCE/DOCKET NUMBER: 053466/0225
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-020-846-68

Query Match 5.9%; Score 7; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||
Db 361 PDKEVLY 367

RESULT 527
US-09-328-352-7646
; Sequence 7646, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7646
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7646

Query Match 5.9%; Score 7; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LYQYQYDE 61
|||
Db 103 LYQYQYDE 109

RESULT 528
US-09-489-039A-8263
; Sequence 8263, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8263
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8263

Query Match 5.9%; Score 7; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GCVLAAL 24
|||
Db 212 GCVLAAL 218

RESULT 529
US-09-252-991A-22682
; Sequence 22682, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22682
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22682

Query Match 5.9%; Score 7; DB 2; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQR 92
|||
Db 142 VLGLLQR 148

RESULT 530
US-09-902-540-11909
; Sequence 11909, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11909
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-11909

Query Match 5.9%; Score 7; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||
Db 465 VLAALAA 471

RESULT 531
US-10-082-894-2
; Sequence 2, Application US/10082894
; Patent No. 6818433
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; APPLICANT: Salmon, Brandy Leigh
; APPLICANT: Bradley, John D.
; TITLE OF INVENTION: NEMATODE PGM-LIKE SEQUENCES
; FILE REFERENCE: 12557-003001
; CURRENT APPLICATION NUMBER: US/10/082,894
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: US 60/271,781
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Meloidogyne incognita PGM
US-10-082-894-2

Query Match 5.9%; Score 7; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 NWQLEA 113
| | | | |
Db 45 NWQLEA 51

RESULT 532
US-10-082-894-4
; Sequence 4, Application US/10082894
; Patent No. 6818433
; GENERAL INFORMATION:
; APPLICANT: Kloeck, Andrew P.
; APPLICANT: Williams, Deryck Jeremy
; APPLICANT: Salmon, Brandy Leigh
; APPLICANT: Bradley, John D.
; TITLE OF INVENTION: NEMATODE PGM-LIKE SEQUENCES
; FILE REFERENCE: 12557-003001
; CURRENT APPLICATION NUMBER: US/10/082,894
; CURRENT FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: US 60/271,781
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Meloidogyne incognita PGM
US-10-082-894-4

Query Match 5.9%; Score 7; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 NWQLEA 113
| | | | |
Db 45 NWQLEA 51

RESULT 533
US-09-252-991A-31967
; Sequence 31967, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31967
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31967

Query Match 5.9%; Score 7; DB 2; Length 808;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 797 VLAALAA 803

RESULT 534
US-09-252-991A-29362
; Sequence 29362, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29362
; LENGTH: 913
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29362

Query Match 5.9%; Score 7; DB 2; Length 913;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 198 VLAALAA 204

RESULT 535
US-10-104-047-2812
; Sequence 2812, Application US/10104047
; Patent No. 6943241
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. 6943241el full length cDNA
; FILE REFERENCE: HI-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2812
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2812


```
Query Match      5.9%; Score 7; DB 2; Length 1036;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 QQQAVIE 101
Db 27 QQQAVIE 33

RESULT 536
US-09-543-681A-4379
; Sequence 4379, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4379
; LENGTH: 1116
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4379

Query Match      5.9%; Score 7; DB 2; Length 1116;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 689 VLAALAA 695

RESULT 537
US-09-605-703B-2408
; Sequence 2408, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habershauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR APPLICATION NUMBER: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 2408
; LENGTH: 1255
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-605-703B-2408

Query Match      5.9%; Score 7; DB 2; Length 1255;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 11
Db 1184 ADLEVT 1190
```

```
RESULT 538
US-10-226-629A-13
; Sequence 13, Application US/10226629A
; Patent No. 6960431
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Therapeutic Compositions and Methods for Treating Viral Infection
; FILE REFERENCE: 5006.01
; CURRENT APPLICATION NUMBER: US/10/226,629A
; CURRENT FILING DATE: 2002-08-22
; PRIOR APPLICATION NUMBER: US 60/314,182
; PRIOR FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 736
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 13
; LENGTH: 2940
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-226-629A-13

Query Match      5.9%; Score 7; DB 2; Length 2940;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 539
US-07-925-695-5
; Sequence 5, Application US/07925695
; Patent No. 5428145
; GENERAL INFORMATION:
; APPLICANT: OKAMOTO, Hiroaki
; APPLICANT: NAKAMURA, Tetsuo
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOME,
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
; TITLE OF INVENTION: DETECTION SYSTEMS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Beveridge, DeGrandi, Weilacher & Young
; STREET: 1850 M Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/925,695
; FILING DATE: 19920807
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 287402/91
; FILING DATE: 09-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 360441/91
; FILING DATE: 05-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Weilacher, Robert G.
; REGISTRATION NUMBER: 20,531
; REFERENCE/DOCKET NUMBER: 06/87-48009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 659-2811
; TELEFAX: (202) 659-1462
; TELEX: WUI 64470
```

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3033 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: linear
US-07-925-695-5

Query Match 5.9%; Score 7; DB 1; Length 3033;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 540

US-08-444-818-458
; Sequence 458, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; TITLE OF INVENTION: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424

; PRIOR APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885

; INFORMATION FOR SEQ ID NO: 458:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-444-818-458

Query Match 5.1%; Score 6; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TSTWVL 16
Db 3 TSTWVL 8

RESULT 541

US-08-146-028-283
; Sequence 283, Application US/08146028

; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-283

Query Match 5.1%; Score 6; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 542

US-08-146-028-289
; Sequence 289, Application US/08146028
; Patent No. 5891640
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; INFORMATION FOR SEQ ID NO: 289:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-146-028-289

Query Match 5.1%; Score 6; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 543

US-08-146-028-294

Sequence 294, Application US/08146028
Patent No. 5891640

GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/146,028

INFORMATION FOR SEQ ID NO: 294:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-146-028-294

Query Match 5.1%; Score 6; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 MEECSQ 67

DB 1 MEECSQ 6

RESULT 544

US-08-146-028-300

Sequence 300, Application US/08146028
Patent No. 5891640

GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/146,028

INFORMATION FOR SEQ ID NO: 300:

SEQUENCE CHARACTERISTICS:

LENGTH: 9 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-146-028-300

Query Match 5.1%; Score 6; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 MEECSQ 67

DB 1 MEECSQ 6

RESULT 545

US-08-146-028-283

Sequence 283, Application US/08723425A
Patent No. 6165730

GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

US-08-723-425A-283

;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/723,425A
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SADOFF, B.J.
;; REGISTRATION NUMBER: 36,663
;; REFERENCE/DOCKET NUMBER: 1487-13
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-816-4000
;; TELEFAX: 703-816-4100
;; INFORMATION FOR SEQ ID NO: 289:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-723-425A-289

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 547
US-08-723-425A-294
; Sequence 294, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 294:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-723-425A-294

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 62 MEECSQ 67
Db 1 MEECSQ 6
RESULT 548
US-08-723-425A-300
; Sequence 300, Application US/08723425A
; Patent No. 6165730
; GENERAL INFORMATION:
; APPLICANT: DELEYS, ROBERT
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
; TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
; NUMBER OF SEQUENCES: 453
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE, P.C.
; STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,425A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: SADOFF, B.J.
; REGISTRATION NUMBER: 36,663
; REFERENCE/DOCKET NUMBER: 1487-13
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 300:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-723-425A-300

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 MEECSQ 67
Db 1 MEECSQ 6

RESULT 549
US-09-112-206-283
; Sequence 283, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-283

Query Match          5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
DB 4 DEMEEC 9

RESULT 550
US-09-112-206-289
; Sequence 289, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-289

Query Match          5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
DB 4 DEMEEC 9

RESULT 551
US-09-112-206-294
; Sequence 294, Application US/09112206
```

```

; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 294:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-294

Query Match          5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 MEECSQ 67
DB 1 MEECSQ 6

RESULT 552
US-09-112-206-300
; Sequence 300, Application US/09112206
; Patent No. 6210903
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/112,206
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/146,028
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 300:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-112-206-300

Query Match          5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      62 MEECSQ 67
      |||||
Db      1 MEECSQ 6

RESULT 553
US-09-311-784A-257
; Sequence 257, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 257
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV NS4 1666 (peptide 24.0075)
US-09-311-784A-257

Query Match      5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLA 23
      |||||
Db      4 GGVLA 9

RESULT 554
US-09-311-784A-450
; Sequence 450, Application US/09311784A
; Patent No. 6534482
; GENERAL INFORMATION:
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE REFERENCE: 39963-20022.01
; CURRENT APPLICATION NUMBER: US/09/311,784A
; CURRENT FILING DATE: 1999-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 450
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Pf EXP1 91 (peptide 1167.09)
US-09-311-784A-450
```

```
Query Match      5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      15 VLLGV 20
      |||||
Db      1 VLLGV 6

RESULT 555
US-09-790-497A-273
; Sequence 273, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 273
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-273

Query Match      5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEEC 65
      |||||
Db      4 DEMEEC 9

RESULT 556
US-09-790-497A-278
; Sequence 278, Application US/09790497A
; Patent No. 6649735
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; FILE REFERENCE: 2752-16
; CURRENT APPLICATION NUMBER: US/09/790,497A
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: 09/576,824
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
US-09-790-497A-278
```

; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 278
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-790-497A-278

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 MEECSQ 67
Db 1 MEECSQ 6

RESULT 557
US-09-576-824A-273
; Sequence 273, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11

; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 273
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-273

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 558
US-09-576-824A-278
; Sequence 278, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11

; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 278
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-576-824A-278

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 62 MEECSQ 67
Db 1 MEECSQ 6

RESULT 559
US-09-680-497-283
; Sequence 283, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 283:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-680-497-283

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 560
US-09-680-497-289
; Sequence 289, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:

```
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 289:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-680-497-289

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 561
US-09-680-497-294
; Sequence 294, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 294:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-680-497-294

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 MEECSQ 67
```

```
Db 1 MEECSQ 6

RESULT 562
US-09-680-497-300
; Sequence 300, Application US/09680497
; Patent No. 6709828
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
; TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
; TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
; TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
; TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
; NUMBER OF SEQUENCES: 453
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/680,497
; FILING DATE: 06-OCT-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/146,028
; FILING DATE: 22-NOV-1993
; INFORMATION FOR SEQ ID NO: 300:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-680-497-300

Query Match 5.1%; Score 6; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.6e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 62 MEECSQ 67
Db 1 MEECSQ 6

RESULT 563
US-08-214-650-33
; Sequence 33, Application US/08214650
; Patent No. 5709995
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; APPLICANT: Cerny, Andreas
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,650
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Silvert, Donald J.
```


REGISTRATION NUMBER: 37552
REFERENCE/DOCKET NUMBER: 61230
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
TELEX: 25-3533
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-214-650-33

Query Match 5.1%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWIV 37
DB 5 GCWIV 10

RESULT 564

US-08-617-929-13
Sequence 13, Application US/08617929
Patent No. 5885771

GENERAL INFORMATION:
APPLICANT: KUMAZAWA, Toshiaki
TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
TITLE OF INVENTION: IMMUNOASSAY
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,929
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/01823
FILING DATE: 28-OCT-1994

PRIOR APPLICATION NUMBER: JP 6/207695
FILING DATE: 31-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5/272864
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 77384/109
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-929-13

Query Match 5.1%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 PAIVPD 51
DB 5 PAIVPD 10

RESULT 565

US-08-617-929-33
Sequence 33, Application US/08617929
Patent No. 5885771

GENERAL INFORMATION:
APPLICANT: KUMAZAWA, Toshiaki
TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
TITLE OF INVENTION: IMMUNOASSAY
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,929
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/01823
FILING DATE: 28-OCT-1994

PRIOR APPLICATION NUMBER: JP 6/207695
FILING DATE: 31-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5/272864
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 77384/109
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-929-33

Query Match 5.1%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
DB 4 DEMEEC 9

RESULT 566

US-08-617-929-38
Sequence 38, Application US/08617929
Patent No. 5885771

;; GENERAL INFORMATION:
;; APPLICANT: KOMAZAWA, Toshiaki
;; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
;; TITLE OF INVENTION: IMMUNOASSAY
;; NUMBER OF SEQUENCES: 42
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 3000 K Street, N.W., Suite 500
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/617,929
;; FILING DATE: 24-APR-1996
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: WO PCT/JP94/01823
;; FILING DATE: 28-OCT-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 6/207695
;; FILING DATE: 31-AUG-1994
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 5/272864
;; FILING DATE: 29-OCT-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SAXE, Bernhard D.
;; REGISTRATION NUMBER: 28,665
;; REFERENCE/DOCKET NUMBER: 77384/109
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202)672-5300
;; TELEX: (202)672-5399
;; TELEFAX: 904136
;; INFORMATION FOR SEQ ID NO: 38:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; US-08-617-929-38

Query Match 5.1%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 567
US-08-802-981-211
; Sequence 211, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/802,981
;; FILING DATE: 20-FEB-1997
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hunter, Tom
;; REGISTRATION NUMBER: 38,498
;; REFERENCE/DOCKET NUMBER: 016865-000300US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 576-0200
;; TELEFAX: (415) 576-0300
;; INFORMATION FOR SEQ ID NO: 211:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 10 amino acids
;; TYPE: amino acid
;; STRANDEDNESS:
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-802-981-211

Query Match 5.1%; Score 6; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 568
US-09-011-961-19
; Sequence 19, Application US/09011961
; Patent No. 6197536
; GENERAL INFORMATION:
; APPLICANT: STEINKUEHLER, Christian
; APPLICANT: PESSI, Antonello
; APPLICANT: BIANCHI, Elisabetta
; APPLICANT: TALIANI, Marina
; APPLICANT: TOMELI, Licia
; APPLICANT: URBANI, Andrea
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: NARJES, Frank
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
; TITLE OF INVENTION: HCV NS3 PROTEASE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,961
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IT96/00163
; FILING DATE: 20-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM95A000573
; FILING DATE: 22-AUG-1995
; ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: STEINKUEHLER=1
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-011-961-19

Query Match 5.1%; Score 6; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 569

US-09-011-961-35
Sequence 35, Application US/09011961
Patent No. 6197536

GENERAL INFORMATION:
APPLICANT: STEINKUEHLER, Christian
APPLICANT: PESSI, Antonello
APPLICANT: BIANCHI, Elisabetta
APPLICANT: TALIANI, Marina
APPLICANT: TOMEI, Licia
APPLICANT: URBANI, Andrea
APPLICANT: DE FRANCESCO, Raffaele
APPLICANT: NARJES, Frank
TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/011,961
FILING DATE: 23-FEB-1998
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IT96/00163
FILING DATE: 20-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM95A000573
FILING DATE: 22-AUG-1995
ATTORNEY/AGENT INFORMATION:

NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: STEINKUEHLER=1
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /note= "Xaa is Ac-Asp"
US-09-011-961-35

Query Match 5.1%; Score 6; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 EMEECs 66
Db 2 EMEECs 7

RESULT 570

US-09-311-784A-435
Sequence 435, Application US/09311784A
Patent No. 6534482

GENERAL INFORMATION:
APPLICANT: Fikes, John D.
APPLICANT: Hermanson, Gary G.
APPLICANT: Sette, Alessandro
APPLICANT: Ishioka, Glenn Y.
APPLICANT: Livingston, Brian
APPLICANT: Chesnut, Robert W.
APPLICANT: Epimmune Inc.
TITLE OF INVENTION: Expression Vectors for Stimulating an
TITLE OF INVENTION: Immune Response and Methods of Using the Same
FILE REFERENCE: 39963-20022.01
CURRENT APPLICATION NUMBER: US/09/311,784A
PRIOR FILING DATE: 1999-05-13
PRIOR APPLICATION NUMBER: US 60/085,751
NUMBER OF SEQ ID NOS: 463
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 435
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: pf EXP1 91 (peptide 1167.19)
US-09-311-784A-435

Query Match 5.1%; Score 6; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 1 VLLGGV 6

RESULT 571

US-09-011-961-13
Sequence 13, Application US/09011961
Patent No. 6197536

GENERAL INFORMATION:
APPLICANT: STEINKUEHLER, Christian
APPLICANT: PESSI, Antonello
APPLICANT: BIANCHI, Elisabetta
APPLICANT: TALIANI, Marina
APPLICANT: TOMEI, Licia
APPLICANT: URBANI, Andrea
APPLICANT: DE FRANCESCO, Raffaele
APPLICANT: NARJES, Frank
TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE

;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
;; STREET: 419 Seventh Street, N.W., Suite 300
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/011,961
;; FILING DATE: 23-FEB-1998
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: PCT/IT96/00163
;; FILING DATE: 20-AUG-1996
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: IT RM95A000573
;; FILING DATE: 22-AUG-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: YUN, Allen C.
;; REGISTRATION NUMBER: 37,971
;; REFERENCE/DOCKET NUMBER: STEINKUHLER-1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 11 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1
;; OTHER INFORMATION: /note="Xaa is Ac-Tyr"
US-09-011-961-13

Query Match 5.1%; Score 6; DB 2; Length 11;
Best Local Similarity 100.0%; Pred.No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMREC 65
Db 5 DEMREC 10
|||||

RESULT 572
US-09-576-824A-528
; Sequence 528, Application US/09576824A
; Patent No. 6667387
; GENERAL INFORMATION:
; APPLICANT: De Leys, Robert
; TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
; TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
; TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
; TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
; TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
; TITLE OF INVENTION: CONTAINING THEM
; FILE REFERENCE: 2752-11
; CURRENT APPLICATION NUMBER: US/09/576,824A
; CURRENT FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 08/723,425
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08

;; PRIOR APPLICATION NUMBER: EP 92400598.6
;; PRIOR FILING DATE: 1992-03-06
;; NUMBER OF SEQ ID NOS: 600
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 528
;; LENGTH: 11
;; TYPE: PRT
;; ORGANISM: Hepatitis C virus
;; FEATURE:
;; NAME/KEY: VARIANT
;; LOCATION: (1)
;; OTHER INFORMATION: Xaa = modified site : when present, represents an
;; amino acid, amino group, or chemically modified
;; OTHER INFORMATION: amino terminus
;; NAME/KEY: VARIANT
;; LOCATION: (11)
;; OTHER INFORMATION: Xaa = modified site : when present, represents an
;; amino acid, OH-group, NH2-group, or a linkage
;; OTHER INFORMATION: involv-ing these two groups
US-09-576-824A-528

Query Match 5.1%; Score 6; DB 2; Length 11;
Best Local Similarity 100.0%; Pred.No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 APYIEQ 74
Db 1 APYIEQ 6
|||||

RESULT 573
US-08-617-929-17
; Sequence 17, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-617-929-17

Query Match 5.1%; Score 6; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 PAIPVD 51
Db 1 PAIPVD 6

RESULT 574
US-08-853-623D-11
; Sequence 11, Application US/08853623D
; Patent No. 5990276
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Mui, Philip
; APPLICANT: Weber, Patricia
; TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
; TITLE OF INVENTION: Virus NS3 Protease
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,623D
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER: US 60/017,470
; FILING DATE: 10-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0595
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; OTHER INFORMATION: The amino acid residues are
; OTHER INFORMATION: preferably D-amino acid residues.
US-08-853-623D-11

Query Match 5.1%; Score 6; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CWIVVG 38
Db 1 CWIVVG 6

RESULT 575
US-09-011-961-12
; Sequence 12, Application US/09011961
; Patent No. 6197536
; GENERAL INFORMATION:
; APPLICANT: STEINKUEHLER, Christian
; APPLICANT: PESSI, Antonello
; APPLICANT: BIANCHI, Elisabetta
; APPLICANT: TALIANI, Marina
; APPLICANT: TOMEI, Licia
; APPLICANT: URBANI, Andrea
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: NARJES, Frank
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,961
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IT96/00163
; FILING DATE: 20-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM95A000573
; FILING DATE: 22-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: STEINKUEHLER-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa is Ac-Tyr"
US-09-011-961-12

Query Match 5.1%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 5 DEMEEC 10

RESULT 576
US-09-403-752A-78
; Sequence 78, Application US/09403752A

; Patent No. 6593132
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: RICIN-LIKE TOXIN VARIANTS FOR TREATMENT OF CANCER,
; FILE REFERENCE: 10447-005
; CURRENT APPLICATION NUMBER: US/09/403,752A
; CURRENT FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: U.S. 60/045,148
; PRIOR FILING DATE: 1997-04-30
; PRIOR APPLICATION NUMBER: U.S. 60/063,715
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Mutant preporicrin linker region for HCV-B, pAP-264
US-09-403-752A-78

Query Match 5.1%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 577
US-09-551-151A-78
; Sequence 78, Application US/09551151A
; Patent No. 6803358
; GENERAL INFORMATION:
; APPLICANT: Borgford, Thor
; TITLE OF INVENTION: Ricin-Like Toxin Variants for Treatment of Cancer,
; FILE REFERENCE: 10447-011
; CURRENT APPLICATION NUMBER: US/09/551,151A
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US 09/403,752
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 78
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant preporicrin linker region for HCV-B, pAP-264
US-09-551-151A-78

Query Match 5.1%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 578
US-09-929-955-28
; Sequence 28, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS3/4A peptide
US-09-929-955-28

Query Match 5.1%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SADLEV 9
Db 1 SADLEV 6

RESULT 579
US-09-930-591-15
; Sequence 15, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis C virus NS3/4A peptide
US-09-930-591-15

Query Match 5.1%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SADLEV 9
Db 1 SADLEV 6

RESULT 580
US-08-439-747A-15
; Sequence 15, Application US/08439747A
; Patent No. 5767233
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Murray, Michael
; APPLICANT: Ramanathan, Lata
; TITLE OF INVENTION: Soluble, Cleavable Substrates of the Hepatitis
; FILE REFERENCE: TRIPEP.23AUS2
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:

ADDRESS: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.3
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,747A
FILING DATE: May 12, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: JB0509
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-298-5061
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
FEATURE:
NAME/KEY: NS4A Active Mutant
US-08-439-747A-15

Query Match 5.1%; Score 6; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CWWIVG 38
DB 1 CWWIVG 6

RESULT 581
US-08-440-409B-15
Sequence 15, Application US/08440409B
Patent No. 5843752
GENERAL INFORMATION:
APPLICANT: Dasamhapatra, Bimal
APPLICANT: Murray, Michael
APPLICANT: Ramanathan, Lata
APPLICANT: Butkiewicz, Nancy
TITLE OF INVENTION: Soluble Active Hepatitis C Virus Protease
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.3
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,409B
FILING DATE: May 12, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: JB0494

TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-298-5061
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
FEATURE:
NAME/KEY: NS4A Active Mutant
US-08-440-409B-15

Query Match 5.1%; Score 6; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CWWIVG 38
DB 1 CWWIVG 6

RESULT 582
US-08-853-623D-21
Sequence 21, Application US/08853623D
Patent No. 5990276
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Mui, Philip
APPLICANT: Weber, Patricia
TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,623D
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,470
FILING DATE: 10-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: NS4A Active Mutant
US-08-853-623D-21

Query Match 5.1%; Score 6; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 44;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CVWVG 38
| | | | |
Db 1 CVWVG 6

RESULT 583

US-08-853-623D-30
; Sequence 30, Application US/08853623D
; Patent No. 5990276
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Mui, Philip
; TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
; TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853.623D
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,470
; FILING DATE: 10-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; REFERENCE/DOCKET NUMBER: JB0595

TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-08-853-623D-30

Query Match 5.1%; Score 6; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CVWVG 38
| | | | |
Db 1 CVWVG 6

RESULT 584

US-09-011-961-11
; Sequence 11, Application US/09011961
; Patent No. 6197536
; GENERAL INFORMATION:
; APPLICANT: STEINKUHLER, Christian
; APPLICANT: PESSI, Antonello
; APPLICANT: BIANCHI, Elisabetta
; APPLICANT: TALIANI, Marina
; APPLICANT: TOMEI, Licia
; APPLICANT: URBANI, Andrea

; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: NARJES, Frank
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,961
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 435

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IT96/00163
; FILING DATE: 20-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM95A000573
; FILING DATE: 22-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.

REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: STEINKUHLER=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:

NAME/KEY: Peptide

LOCATION: 1

OTHER INFORMATION: /note= "Xaa is Ac-Tyr"

US-09-011-961-11

Query Match 5.1%; Score 6; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
| | | | |
Db 5 DEMEEC 10

RESULT 585

US-09-719-261B-9
; Sequence 9, Application US/09719261B
; Patent No. 6867284
; GENERAL INFORMATION:
; APPLICANT: Istituto Di Ricerche Di Biologia Molecolare P. Angeletti SpA
; APPLICANT: Matassa, Victor
; APPLICANT: Narjes, Frank
; APPLICANT: Koehler, Konrad
; APPLICANT: Ontoria, Jesus
; APPLICANT: Poma, Marco
; TITLE OF INVENTION: Peptide inhibitors of hepatitis C virus NS3 protease
; FILE REFERENCE: KMN/FP5780044
; CURRENT APPLICATION NUMBER: US/09/719,261B
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB99/01824

;; PRIOR FILING DATE: 1999-06-09
;; PRIOR APPLICATION NUMBER: GB 9812523.0
;; PRIOR FILING DATE: 1998-06-10
;; NUMBER OF SEQ ID NOS: 13
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 9
;; LENGTH: 13
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;; OTHER INFORMATION: sequence
US-09-719-261B-9

Query Match 5.1%; Score 6; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 586
US-08-232-453A-7
; Sequence 7, Application US/08232453A
; Patent No. 5589568
; GENERAL INFORMATION:
; APPLICANT: HIGASHIJIMA, TSUTOMU
; APPLICANT: ROSS, ELLIOTT M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: MODULATING G PROTEIN ACTION
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,453A
; FILING DATE: APRIL 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,319
; FILING DATE: AUGUST 21, 1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:253/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-232-453A-7

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26

Db 6 LAALAA 11

RESULT 587
US-08-232-453A-17
; Sequence 17, Application US/08232453A
; Patent No. 5589568
; GENERAL INFORMATION:
; APPLICANT: HIGASHIJIMA, TSUTOMU
; APPLICANT: ROSS, ELLIOTT M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: MODULATING G PROTEIN ACTION
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,453A
; FILING DATE: APRIL 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,319
; FILING DATE: AUGUST 21, 1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:253/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-232-453A-17

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 3 LAALAA 8

RESULT 588
US-08-232-453A-25
; Sequence 25, Application US/08232453A
; Patent No. 5589568
; GENERAL INFORMATION:
; APPLICANT: HIGASHIJIMA, TSUTOMU
; APPLICANT: ROSS, ELLIOTT M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: MODULATING G PROTEIN ACTION
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TX

```
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: IBM PC COMPATIBLE
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.453A
; FILING DATE: APRIL 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,319
; FILING DATE: AUGUST 21, 1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:253/PAR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-232-453A-25

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 3 LAALAA 8

RESULT 589
US-08-232-453A-26
; Sequence 26, Application US/08232453A
; Patent No. 5589568
; GENERAL INFORMATION:
; APPLICANT: HIGASHIJIMA, TSUTOMU
; APPLICANT: ROSS, ELLIOTT M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; TITLE OF INVENTION: MODULATING G PROTEIN ACTION
; NUMBER OF SEQUENCES: 71
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ARNOLD, WHITE & DURKEE
; STREET: P.O. BOX 4433
; CITY: HOUSTON
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; OPERATING SYSTEM: IBM PC COMPATIBLE
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.453A
; FILING DATE: APRIL 22, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/748,319
; FILING DATE: AUGUST 21, 1991
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: PARKER, DAVID L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: UTSD:253/PAR
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (512) 474-7577
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-232-453A-26

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 6 LAALAA 11

RESULT 590
US-08-617-929-16
; Sequence 16, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KOMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; TITLE OF INVENTION: IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-617-929-16

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 46 PAIVPD 51
Db 2 PAIVPD 7

RESULT 591

US-08-617-929-36
; Sequence 36, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; TITLE OF INVENTION: IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; APPLICATION DATA:
; JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-617-929-36

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 592

US-08-617-929-41
; Sequence 41, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; TITLE OF INVENTION: IMMUNOASSAY

; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; APPLICATION DATA:
; JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-617-929-41

Query Match 5.1%; Score 6; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 1 DEMEEC 6

RESULT 593

US-09-011-961-10
; Sequence 10, Application US/09011961
; Patent No. 6197536
; GENERAL INFORMATION:
; APPLICANT: STEINKUEHLER, Christian
; APPLICANT: PESSI, Antonello
; APPLICANT: BIANCHI, Elisabetta
; APPLICANT: TALIANI, Marina
; APPLICANT: TOMEI, Licia
; APPLICANT: URBANI, Andrea
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: NARJES, Frank
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
; TITLE OF INVENTION: HCV NS3 PROTEASE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.

;; COUNTRY: USA
;; ZIP: 20004
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA: US/09/011.961
;; FILING DATE: 23-FEB-1998
;; CLASSIFICATION: 435
;; PRIOR APPLICATION NUMBER: PCT/IT96/00163
;; FILING DATE: 20-AUG-1996
;; APPLICATION NUMBER: IT RM95A000573
;; FILING DATE: 22-AUG-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: YUN, Allen C.
;; REGISTRATION NUMBER: 37,971
;; REFERENCE/DOCKET NUMBER: STEINKUHLER=1
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 10:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 14 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1
;; OTHER INFORMATION: /note= "Xaa is Ac-Tyr"
US-09-011-961-10

Query Match 5.1%; Score 6; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 5 DEMEEC 10

RESULT 594
US-09-011-961-9
; Sequence 9, Application US/09011961
; Patent No. 6197536
; GENERAL INFORMATION:
; APPLICANT: STEINKUHLER, Christian
; APPLICANT: PESSI, Antonello
; APPLICANT: BIANCHI, Elisabetta
; APPLICANT: TALIANI, Marina
; APPLICANT: TOMEI, Licia
; APPLICANT: URBANI, Andrea
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: NARJES, Frank
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA: US/09/011.961
;; FILING DATE: 23-FEB-1998
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA: PCT/IT96/00163
;; FILING DATE: 20-AUG-1996
;; APPLICATION NUMBER: IT RM95A000573
;; FILING DATE: 22-AUG-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: YUN, Allen C.
;; REGISTRATION NUMBER: 37,971
;; REFERENCE/DOCKET NUMBER: STEINKUHLER=1
;; TELEPHONE: 202-628-5197
;; TELEFAX: 202-737-3528
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 15 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Peptide
;; LOCATION: 1
;; OTHER INFORMATION: /note= "Xaa is Ac-Tyr"
US-09-011-961-9

Query Match 5.1%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 5 DEMEEC 10

RESULT 595
US-09-009-953-27
; Sequence 27, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly
; REACTIVE DR RESTRICTED EPITOPES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009.953
; FILING DATE: 21-Jan-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,713
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver

REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-011520US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-009-953-27

Query Match 5.1%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
DB 10 VLLGGV 15

RESULT 596
US-09-009-953-55
; Sequence 55, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly
; REACTIVE DR Restricted Epitopes
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,953
; FILING DATE: 21-Jan-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,713
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-011520US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-009-953-55

Query Match 5.1%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
DB 10 VLLGGV 15

RESULT 597
US-09-009-953-70
; Sequence 70, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly
; REACTIVE DR Restricted Epitopes
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,953
; FILING DATE: 21-Jan-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,713
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-011520US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 70:
US-09-009-953-70

Query Match 5.1%; Score 6; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
DB 7 VLLGGV 12

RESULT 598
US-09-009-953-81
; Sequence 81, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly

US-09-011-961-8
; Sequence 8, Application US/09011961
; Patent No. 6197536
; GENERAL INFORMATION:
; APPLICANT: STEINKUEHLER, Christian
; APPLICANT: PESSI, Antonello
; APPLICANT: BIANCHI, Elisabetta
; APPLICANT: TALIANI, Marina
; APPLICANT: TOMBI, Licia
; APPLICANT: URBANI, Andrea
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: NARJES, Frank
; TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
; TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
; TITLE OF INVENTION: HCV NS3 PROTEASE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/011,961
; FILING DATE: 23-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/IT96/00163
; FILING DATE: 20-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IT RM95A000573
; FILING DATE: 22-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C.
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: STINKUEHLER=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1
; OTHER INFORMATION: /note= "Xaa is Ac-Tyr"
US-09-011-961-8

Query Match 5.1%; Score 6; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 5 DEMEEC 10
|||||

RESULT 602
US-08-537-802-32
; Sequence 32, Application US/08537802
; Patent No. 6881821
; GENERAL INFORMATION:
; APPLICANT:

; TITLE OF INVENTION: HEPATITIS-C VIRUS TYPE 4, 5 & 6
; NUMBER OF SEQUENCES: 50
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,802
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/00957
; FILING DATE:
; APPLICATION NUMBER: GB 9309237.7
; FILING DATE: 05-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9400263.1
; FILING DATE: 07-JAN-1994
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-537-802-32

Query Match 5.1%; Score 6; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQ 58
Db 9 EVLYQQ 14
|||||

RESULT 603
US-08-537-802-34
; Sequence 34, Application US/08537802
; Patent No. 6881821
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS-C VIRUS TYPE 4, 5 & 6
; NUMBER OF SEQUENCES: 50
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,802
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB94/00957
; FILING DATE:
; APPLICATION NUMBER: GB 9309237.7
; FILING DATE: 05-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9400263.1
; FILING DATE: 07-JAN-1994
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-537-802-34

Query Match 5.1%; Score 6; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQ 58
|||||

Db 9 EVLYQQ 14

RESULT 604
US-08-802-981-94
; Sequence 94, Application US/08802981
; Patent No. 6037137
; GENERAL INFORMATION:
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly S.
; TITLE OF INVENTION: Compositions for the Detection of Enzyme
; TITLE OF INVENTION: Activity in Biological Samples and Methods of Use Thereof
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/802,981
; FILING DATE: 20-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 016865-0003000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 94:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "Aib"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4
; OTHER INFORMATION: /product= "Acp"
US-08-802-981-94

Query Match 5.1%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
|||||
Db 6 DEMEEC 11

RESULT 605
US-09-747-287A-137
; Sequence 137, Application US/09747287A
; Patent No. 6893868
; GENERAL INFORMATION:
; APPLICANT: KOMORIYA, AKIRA
; APPLICANT: PACKARD, BEVERLY S.
; TITLE OF INVENTION: HOMO-DOUBLY LABELED COMPOSITIONS FOR THE DETECTION OF ENZYME
; TITLE OF INVENTION: ACTIVITY IN BIOLOGICAL SAMPLES
; FILE REFERENCE: 300-948600US
; CURRENT APPLICATION NUMBER: US/09/747,287A

Query Match 5.1%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
|||||
Db 6 DEMEEC 11

RESULT 606
US-09-394-019C-90
; Sequence 90, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncoimmunin, Inc.
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903820US
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 90
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide substrate
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (3)..(3)
; OTHER INFORMATION: X is Aib
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (3)..(4)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (4)..(4)
; OTHER INFORMATION: X is epsilon aminocaproic acid
US-09-394-019C-90

Query Match 5.1%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
|||||
Db 6 DEMEEC 11

Query Match 5.1%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 6 DEMEEC 11

RESULT 607
US-09-394-019C-335
; Sequence 335, Application US/09394019C
; Patent No. 6936687
; GENERAL INFORMATION:
; APPLICANT: Oncomeriva, Inc.
; APPLICANT: Komoriya, Akira
; APPLICANT: Packard, Beverly
; TITLE OF INVENTION: COMPOSITIONS FOR THE DETECTION OF ENZYME ACTIVITY IN BIOLOGICAL
; FILE REFERENCE: 300-903820US
; CURRENT APPLICATION NUMBER: US/09/394,019C
; CURRENT FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: PCT/US98/00300
; PRIOR FILING DATE: 1998-02-20
; PRIOR APPLICATION NUMBER: US 08/802,981
; PRIOR FILING DATE: 1997-02-20
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 335
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic peptide. Chemically synthesized protease substrate.
; NAME/KEY: misc_feature
; LOCATION: (3)..(3)
; OTHER INFORMATION: Xaa is alpha-aminoisobutyric acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4)..(4)
; OTHER INFORMATION: Xaa is epsilon-aminocaproic acid
US-09-394-019C-335

Query Match 5.1%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 62; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0

QY 60 DEMEEC 65
Db 6 DEMEEC 11

RESULT 608
US-08-617-929-1
; Sequence 1, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; TITLE OF INVENTION: IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; FILING DATE: 24-APR-1996

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-617-929-1

Query Match 5.1%; Score 6; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 PAIVPD 51
Db 5 PAIVPD 10

RESULT 609
US-08-617-929-5
; Sequence 5, Application US/08617929
; Patent No. 5885771
; GENERAL INFORMATION:
; APPLICANT: KUMAZAWA, Toshiaki
; TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
; TITLE OF INVENTION: IMMUNOASSAY
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/617,929
; FILING DATE: 24-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP94/01823
; FILING DATE: 28-OCT-1994
; APPLICATION NUMBER: JP 6/207695
; FILING DATE: 31-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5/272864
; FILING DATE: 29-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: SAXE, Bernhard D.
; REGISTRATION NUMBER: 28,665
; REFERENCE/DOCKET NUMBER: 77384/109

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-929-5

Query Match 5.1%; Score 6; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 610
US-08-617-929-6
Sequence 6, Application US/08617929
Patent No. 5885771
GENERAL INFORMATION:
APPLICANT: KUMAZAWA, Toshiaki
TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
TITLE OF INVENTION: IMMUNOASSAY
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,929
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/01823
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6/207695
FILING DATE: 31-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5/272864
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 77384/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-929-6

Query Match 5.1%; Score 6; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEEC 65
Db 4 DEMEEC 9

RESULT 611
US-08-853-623D-8
Sequence 8, Application US/08853623D
Patent No. 5990276
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Mui, Philip
TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
TITLE OF INVENTION: Virus NS3 Protease
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,623D
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,470
FILING DATE: 10-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY:
OTHER INFORMATION: Xaa is a lysine wherein
the e amino group of which forms a peptide bond with the
carboxyl group of the cysteine residue at position 8 and
OTHER INFORMATION: the carboxyl group of the lysine residue forms a peptide
OTHER INFORMATION: bond with an amino group of another lysine residue (not
OTHER INFORMATION: shown), preferably the amino acid residues at positions
OTHER INFORMATION: 8 - 20 are D- amino acid residues.
US-08-853-623D-8

Query Match 5.1%; Score 6; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CWWIVG 38
Db 8 CWWIVG 13

RESULT 612
US-09-011-961-7

Sequence 7, Application US/09011961
Patent No. 6197536
GENERAL INFORMATION:
APPLICANT: STEINKUEHLER, Christian
APPLICANT: PESSI, Antonello
APPLICANT: BIANCHI, Elisabetta
APPLICANT: TALIANI, Marina
APPLICANT: TOMEI, Lucia
APPLICANT: URBANI, Andrea
APPLICANT: DE FRANCESCO, Raffaele
APPLICANT: NARJES, Frank
TITLE OF INVENTION: METHODOLOGY TO PRODUCE, AND PURIFY AND
TITLE OF INVENTION: ASSAY POLYPEPTIDES WITH THE PROTEOLYTIC ACTIVITY OF THE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/011,961
FILING DATE: 23-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/IT96/00163
FILING DATE: 20-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IT RM95A000573
FILING DATE: 22-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: YUN, Allen C.
REGISTRATION NUMBER: 37,971
REFERENCE/DOCKET NUMBER: STEINKUEHLER=1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1
OTHER INFORMATION: /note= "Xaa is Pmoc-Tyr"
US-09-011-961-7

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 5 DEMEEC 10

RESULT 613
US-09-208-966-48
Sequence 48, Application US/09208966
Patent No. 6221355
GENERAL INFORMATION:
APPLICANT: Dowdy, Steven F.
TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF

FILE REFERENCE: 48881/1742
CURRENT APPLICATION NUMBER: US/09/208,966
CURRENT FILING DATE: 1998-12-10
EARLIER APPLICATION NUMBER: 60/082,402
EARLIER FILING DATE: 1998-04-20
EARLIER APPLICATION NUMBER: 60/069,012
EARLIER FILING DATE: 1997-12-10
NUMBER OF SEQ ID NOS: 57
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 48
LENGTH: 20
TYPE: PRT
ORGANISM: human
US-09-208-966-48

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 5 DEMEEC 10

RESULT 614
US-09-775-052A-48
Sequence 48, Application US/09775052A
Patent No. 6645501
GENERAL INFORMATION:
APPLICANT: Dowdy, Steven F.
TITLE OF INVENTION: ANTI-PATHOGEN SYSTEM AND METHODS OF USE THEREOF
FILE REFERENCE: 48881/1742
CURRENT APPLICATION NUMBER: US/09/775,052A
CURRENT FILING DATE: 2001-12-05
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/208,966
PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-10
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/069,012
PRIOR FILING DATE: EARLIER FILING DATE: 1997-12-10
NUMBER OF SEQ ID NOS: 57
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 48
LENGTH: 20
TYPE: PRT
ORGANISM: human
US-09-775-052A-48

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 5 DEMEEC 10

RESULT 615
US-09-790-497A-94
Sequence 94, Application US/09790497A
Patent No. 6649735
GENERAL INFORMATION:
APPLICANT: De Leys, Robert
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
A PROCESS FOR DETERMINATION OF ANTIBODIES OF
TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
CONTAINING THEM
FILE REFERENCE: 2752-16
CURRENT APPLICATION NUMBER: US/09/790,497A
CURRENT FILING DATE: 2001-02-23
PRIOR APPLICATION NUMBER: 09/576,824
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 08/723,425

```
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: 09/146,028
; PRIOR FILING DATE: 1993-11-22
; PRIOR APPLICATION NUMBER: PCT/EP93/00517
; PRIOR FILING DATE: 1993-03-08
; PRIOR APPLICATION NUMBER: EP 92400598.6
; PRIOR FILING DATE: 1992-03-06
; NUMBER OF SEQ ID NOS: 600
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 94
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; US-09-790-497A-94

Query Match          5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DMEEC 65
Db 1 DMEEC 6

RESULT 616
US-09-878-281A-103
; Sequence 103, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Innogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis, and therapy
; FILE REFERENCE: 35
; CURRENT FILING DATE: 2001-06-12
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 103
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-09-878-281A-103

Query Match          5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 EVLYQQ 58
Db 12 EVLYQQ 17

RESULT 617
US-08-853-623D-1
; Sequence 1, Application US/08853623D
; Patent No. 5990276
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Mui, Philip
; APPLICANT: Weber, Patricia
; TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,623D
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,470
; FILING DATE: 10-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
; TELEFAX: (908)298-5388
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY:
; OTHER INFORMATION: Xaa is aminocaproic acid (Acp).
; US-08-853-623D-1

Query Match          5.1%; Score 6; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CVVIVG 38
Db 9 CVVIVG 14

RESULT 618
US-08-853-623D-2
; Sequence 2, Application US/08853623D
; Patent No. 5990276
; GENERAL INFORMATION:
; APPLICANT: Zhang, Rumin
; APPLICANT: Mui, Philip
; APPLICANT: Weber, Patricia
; TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,623D
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/017,470
; FILING DATE: 10-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McLaughlin, Jaye P.
; REGISTRATION NUMBER: 41,211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)298-5056
```

TELEFAX: (908)298-5388
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY:
OTHER INFORMATION: Xaa is aminocaproic acid (Acp).
US-08-853-623D-2

Query Match 5.1%; Score 6; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CVWVIG 38
| | | | |
DB 8 CVWVIG 13

RESULT 619
US-08-853-623D-6
Sequence 6, Application US/08853623D
Patent No. 5990276
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Mui, Philip
APPLICANT: Weber, Patricia
TITLE OF INVENTION: Synthetic Inhibitors of Hepatitis C
TITLE OF INVENTION: Virus NS3 Protease
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,623D
FILING DATE: 09-MAY-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/017,470
FILING DATE: 10-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: McLaughlin, Jaye P.
REGISTRATION NUMBER: 41,211
REFERENCE/DOCKET NUMBER: JB0595
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)298-5056
TELEFAX: (908)298-5388

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY:

OTHER INFORMATION: Xaa is aminocaproic acid (Acp).
OTHER INFORMATION: The lysine residue at position 8 has a peptide bond between
OTHER INFORMATION: the carboxyl group of Acp and the amino group of the lysine
OTHER INFORMATION: and the amino group of the lysine at position 8 forms a
OTHER INFORMATION: peptide bond with the carboxyl group of the cysteine residue

OTHER INFORMATION: at position 9 and the amino acid residues at positions 9-21
OTHER INFORMATION: are preferably D-amino acid residues;
US-08-853-623D-6

Query Match 5.1%; Score 6; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CVWVIG 38
| | | | |
DB 9 CVWVIG 14

RESULT 620
US-08-146-028-94
Sequence 94, Application US/08146028
Patent No. 5891640
GENERAL INFORMATION:
APPLICANT:

TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,028
INFORMATION FOR SEQ ID NO: 94:

SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV type 2
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-146-028-94

Query Match 5.1%; Score 6; DB 1; Length 22;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
| | | | |
DB 2 DEMEEC 7

RESULT 621
US-08-723-425A-94
Sequence 94, Application US/08723425A
Patent No. 6165730
GENERAL INFORMATION:

APPLICANT: DELEYS, ROBERT
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF
PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES AND THEIR USE IN A PROCESS FOR DETERMINATION OF
TITLE OF INVENTION: ANTIBODIES OR BIOTINYLATED PEPTIDES CORRESPONDING ...
NUMBER OF SEQUENCES: 453
CORRESPONDENCE ADDRESS:

ADDRESSEE: NIXON & VANDERHYZE, P.C.
STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
CITY: Arlington

STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,425A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SADOFF, B.J.
REGISTRATION NUMBER: 36,663
REFERENCE/DOCKET NUMBER: 1487-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV type 2
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-08-723-425A-94

Query Match 5.1%; Score 6; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 2 DEMEEC 7

RESULT 622
US-09-112-206-94
Sequence 94, Application US/09112206
Patent No. 6210903
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
TITLE OF INVENTION: CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/112,206
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/146,028
FILING DATE:
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV type 2
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-09-112-206-94

Query Match 5.1%; Score 6; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 2 DEMEEC 7

RESULT 623
US-09-576-824A-94
Sequence 94, Application US/09576824A
Patent No. 6667387
GENERAL INFORMATION:
APPLICANT: De Leys, Robert
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES CORRESPONDING
TITLE OF INVENTION: TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR USE IN
TITLE OF INVENTION: A PROCESS FOR DETERMINATION OF ANTIBODIES OF
TITLE OF INVENTION: BIOTINYLATED PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT
TITLE OF INVENTION: EPITOPES, A PROCESS FOR PREPARING THEM AND COMPOSITIONS
FILE REFERENCE: 2752-11
CURRENT APPLICATION NUMBER: US/09/576,824A
CURRENT FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 08/723,425
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: 09/146,028
PRIOR FILING DATE: 1993-11-22
PRIOR APPLICATION NUMBER: PCT/EP93/00517
PRIOR FILING DATE: 1993-03-08
PRIOR APPLICATION NUMBER: EP 92400598.6
PRIOR FILING DATE: 1992-03-06
NUMBER OF SEQ ID NOS: 600
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 94
LENGTH: 22
TYPE: PRT
ORGANISM: Hepatitis C virus
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: modified site
NAME/KEY: VARIANT
LOCATION: (22)
OTHER INFORMATION: modified site
US-09-576-824A-94

Query Match 5.1%; Score 6; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 2 DEMEEC 7

RESULT 624
US-09-680-497-94
Sequence 94, Application US/09680497
Patent No. 6709828
GENERAL INFORMATION:

APPLICANT:
TITLE OF INVENTION: PROCESS FOR THE DETERMINATION OF PEPTIDES
CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES AND THEIR
TITLE OF INVENTION: IN A PROCESS FOR DETERMINATION OF ANTIBODIES OR BIOTINYLATED
TITLE OF INVENTION: PEPTIDES CORRESPONDING TO IMMUNOLOGICALLY IMPORTANT EPITOPES,
TITLE OF INVENTION: PROCESS FOR PREPARING THEM AND COMPOSITIONS CONTAINING THEM
NUMBER OF SEQUENCES: 453
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/680,497
FILING DATE: 06-OCT-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/146,028
FILING DATE: 22-NOV-1993
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: HCV type 2
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
FEATURE:
NAME/KEY: Modified-site
LOCATION: 22
US-09-680-497-94

Query Match 5.1%; Score 6; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 2 DEMEEC 7

RESULT 625
US-08-484-635-127
Sequence 127, Application US/08484635
Patent No 5773569
GENERAL INFORMATION:
APPLICANT: Wrighton, Nicholas C.
APPLICANT: Dower, William J.
APPLICANT: Chang, Ray S.
APPLICANT: Kashyap, Arun K.
APPLICANT: Jolliffe, Linda K.
APPLICANT: Johnson, Dana
APPLICANT: Mulcahy, Linda
TITLE OF INVENTION: Compounds and Peptides That Bind to the
Erythropoietin Receptor
NUMBER OF SEQUENCES: 259
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,635

FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,940
FILING DATE: 19-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 16528A-43-1-1
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 127:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-484-635-127
Query Match 5.1%; Score 6; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 24 LAAYCL 29
Db 4 LAAYCL 9
RESULT 626
US-08-484-631-127
Sequence 127, Application US/08484631
Patent No 5830851
GENERAL INFORMATION:
APPLICANT: Wrighton, Nicholas C.
APPLICANT: Dower, William J.
APPLICANT: Chang, Ray S.
APPLICANT: Kashyap, Arun K.
APPLICANT: Jolliffe, Linda K.
APPLICANT: Johnson, Dana
APPLICANT: Mulcahy, Linda
TITLE OF INVENTION: Compounds and Peptides That Bind to the
Erythropoietin Receptor
NUMBER OF SEQUENCES: 259
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Street Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,631
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/155,940
FILING DATE: 19-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Garrett-Wackowski, Eugenia
REGISTRATION NUMBER: 37,330
REFERENCE/DOCKET NUMBER: 16528A-43-1-2
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 127:
SEQUENCE CHARACTERISTICS:

; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-631-127

Query Match 5.1%; Score 6; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 LAAYCL 29
Db 4 LAAYCL 9

RESULT 627
US-08-827-570-127
; Sequence 127, Application US/08827570
; Patent No. 5986047
; GENERAL INFORMATION:
; APPLICANT: Wrighton, Nicholas C.
; APPLICANT: Dower, William J.
; APPLICANT: Chang, Ray S.
; APPLICANT: Kashyap, Arun K.
; APPLICANT: Jolliffe, Linda K.
; APPLICANT: Johnson, Dana
; APPLICANT: Mulcahy, Linda
; TITLE OF INVENTION: Compounds and Peptides That Bind to the
; TITLE OF INVENTION: Erythropoietin Receptor
; NUMBER OF SEQUENCES: 259
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: One Market Plaza, Steuart Street Tower
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105-1492
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827.570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/484,635
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/155,940
; FILING DATE: 19-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Garrett-Wackowski, Eugenia
; REGISTRATION NUMBER: 37,330
; REFERENCE/DOCKET NUMBER: 16528A-43-1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 127:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-827-570-127

Query Match 5.1%; Score 6; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 LAAYCL 29

Db 4 LAAYCL 9

RESULT 628
US-09-929-955-40
; Sequence 40, Application US/09929955
; Patent No. 6858590
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR FILING DATE: 09/705,547
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-929-955-40

Query Match 5.1%; Score 6; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 TSTWVL 16
Db 15 TSTWVL 20

RESULT 629
US-09-930-591-25
; Sequence 25, Application US/09930591
; Patent No. 6960569
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: A HEPATITIS C VIRUS NON-STRUCTURAL
; TITLE OF INVENTION: NS3/4A FUSION GENE
; FILE REFERENCE: TRIPEP.028AUS
; CURRENT APPLICATION NUMBER: US/09/930,591
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutant Hepatitis C virus NS3/4A peptide
US-09-930-591-25

Query Match 5.1%; Score 6; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 TSTWVL 16


```
Db      15 TSTWVL 20
|||||
Query Match      5.1%; Score 6; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GVLAAAL 24
Db      13 GVLAAAL 18
|||||

RESULT 630
US-09-348-578-1
; Sequence 1, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; EARLIER FILING DATE: 1999-07-07
; EARLIER FILING DATE: 1999-07-07
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(26)
; OTHER INFORMATION: Description of Artificial Sequence:OppA secretion signal
US-09-348-578-1

Query Match      5.1%; Score 6; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GVLAAAL 24
Db      13 GVLAAAL 18
|||||

RESULT 631
US-09-699-684-1
; Sequence 1, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(26)
; OTHER INFORMATION: Description of Artificial Sequence:OppA secretion signal
US-09-699-684-1

Query Match      5.1%; Score 6; DB 2; Length 26;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      19 GVLAAAL 24
Db      13 GVLAAAL 18
|||||

RESULT 632
US-09-348-578-2
; Sequence 2, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; EARLIER FILING DATE: 1999-07-07
; EARLIER FILING DATE: 1999-07-07
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; LOCATION: (1)..(27)
US-09-348-578-2

Query Match      5.1%; Score 6; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GVLAAAL 24
Db      14 GVLAAAL 19
|||||

RESULT 633
US-09-348-578-10
; Sequence 10, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; EARLIER FILING DATE: 1999-07-07
; EARLIER FILING DATE: 1998-07-08
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(27)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-348-578-10
```

Query Match 5.1%; Score 6; DB 2; Length 27;
Best Local Similarity 100.0%; Pred.No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
| | | | |
Db 14 GVLAAAL 19

RESULT 634

US-09-699-684-2
; Sequence 2, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(27)
US-09-699-684-2

Query Match 5.1%; Score 6; DB 2; Length 27;
Best Local Similarity 100.0%; Pred.No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
| | | | |
Db 14 GVLAAAL 19

RESULT 635

US-09-699-684-10
; Sequence 10, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(27)

; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-10

Query Match 5.1%; Score 6; DB 2; Length 27;
Best Local Similarity 100.0%; Pred.No. 86;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
| | | | |
Db 14 GVLAAAL 19

RESULT 636

US-09-348-578-3
; Sequence 3, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(28)
US-09-348-578-3

Query Match 5.1%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred.No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
| | | | |
Db 15 GVLAAAL 20

RESULT 637

US-09-348-578-11
; Sequence 11, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(28)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-11

Query Match 5.1%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 15 GVLAAAL 20

RESULT 638

US-09-348-578-19
; Sequence 19, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(28)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-19

Query Match 5.1%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 15 GVLAAAL 20

RESULT 639

US-09-699-684-3
; Sequence 3, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07

; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
; NAME/KEY: SIGNAL
; LOCATION: (1)..(28)
US-09-699-684-3

Query Match 5.1%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 15 GVLAAAL 20

RESULT 640

US-09-699-684-11
; Sequence 11, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(28)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-11

Query Match 5.1%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 15 GVLAAAL 20

RESULT 641

US-09-699-684-19
; Sequence 19, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684

```
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(28)
; OTHER INFORMATION: Description of Artificial Sequence:Modified Oppa secretion
; OTHER INFORMATION: signal
US-09-699-684-19

Query Match          5.1%; Score 6; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 15 GVLAAL 20

RESULT 642
US-08-336-553A-18
; Sequence 18, Application US/08336553A
; Patent No. 6054264
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; APPLICANT: KUO, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; TITLE OF INVENTION: REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-336-553A-18

Query Match          5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 60 DEMEEC 65
Db 17 DEMEEC 22

RESULT 643
US-09-348-578-4
; Sequence 4, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified Oppa secretion
; OTHER INFORMATION: signal
; NAME/KEY: SIGNAL
; LOCATION: (1)..(29)
US-09-348-578-4

Query Match          5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 16 GVLAAL 21

RESULT 644
US-09-348-578-12
; Sequence 12, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(29)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified Oppa secretion
; OTHER INFORMATION: signal
```

US-09-348-578-12

Query Match 5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 16 GVLAAL 21

RESULT 645

US-09-348-578-20
; Sequence 20, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(29)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Modified OppA secretion
US-09-348-578-20

Query Match 5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 16 GVLAAL 21

RESULT 646

US-08-439-157-18
; Sequence 18, Application US/08439157
; Patent No. 6416944
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; Kuo, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/437,895
; FILING DATE: 09-NO. 6416946-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141

APPLICATION NUMBER: US/08/439,157
FILING DATE: 11-MAY-1995
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553A
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141

INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-439-157-18

Query Match 5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
Db 17 DEMEEC 22

RESULT 647

US-09-437-895-18
; Sequence 18, Application US/09437895
; Patent No. 6416946
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; Kuo, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/437,895
; FILING DATE: 09-NO. 6416946-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141

```
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 29 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-437-895-18

Query Match          5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 17 DEMEEC 22

RESULT 648
US-09-699-684-4
; Sequence 4, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(29)
US-09-699-684-4

Query Match          5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 16 GVLAAL 21

RESULT 649
US-09-699-684-12
; Sequence 12, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(29)
US-09-699-684-12

Query Match          5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 16 GVLAAL 21

RESULT 650
US-09-699-684-20
; Sequence 20, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/348,578
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(29)
US-09-699-684-20

Query Match          5.1%; Score 6; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 16 GVLAAL 21

RESULT 651
US-08-336-553A-20
; Sequence 20, Application US/08336553A
; Patent No. 6054264
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; APPLICANT: KIO, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; TITLE OF INVENTION: REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
```

STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHNHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-336-553A-20

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMERC 65
Db 18 DEMERC 23

RESULT 652
US-09-348-578-5
; Sequence 5, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(30)
US-09-348-578-5

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 17 GVLAAL 22

RESULT 653
US-09-348-578-13
; Sequence 13, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(30)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-348-578-13

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 17 GVLAAL 22

RESULT 654
US-09-348-578-21
; Sequence 21, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(30)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-348-578-21

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-348-578-21

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 17 GVLAAL 22

RESULT 655

US-08-439-157-20
; Sequence 20, Application US/08439157
; Patent No. 6416944
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; Kuo, George
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; REAGENTS FOR USE THEREIN

NUMBER OF SEQUENCES: 75

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/439,157

FILING DATE: 11-May-1995

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/336,553A

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/060,400

FILING DATE: 10-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: LEHNHARDT, SUSAN K.

REGISTRATION NUMBER: 33,943

REFERENCE/DOCKET NUMBER: 22300-20947.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-08-439-157-20

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 18 DEMEEC 23

RESULT 656

US-09-437-895-20
; Sequence 20, Application US/09437895
; Patent No. 6416946
; GENERAL INFORMATION:

APPLICANT: CHIEN, DAVID Y.

Kuo, George

TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
REAGENTS FOR USE THEREIN

NUMBER OF SEQUENCES: 75

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1018

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/437,895

FILING DATE: 09-No. 6416946-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/336,553

FILING DATE: <Unknown>

APPLICATION NUMBER: US 08/060,400

FILING DATE: 10-MAY-1993

ATTORNEY/AGENT INFORMATION:

NAME: LEHNHARDT, SUSAN K.

REGISTRATION NUMBER: 33,943

REFERENCE/DOCKET NUMBER: 22300-20947.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 813-5600

TELEFAX: (415) 494-0792

TELEX: 706141

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

LENGTH: 30 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 20:

US-09-437-895-20

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 18 DEMEEC 23

RESULT 657

US-09-699-684-5
; Sequence 5, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:

APPLICANT: HONJO, Masaru

APPLICANT: NAITOH, Naokazu

APPLICANT: UCHIDA, Hiroshi

APPLICANT: MOCHIZUKI, Daisuke

APPLICANT: MATSUMOTO, Kazuya

TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE

FILE REFERENCE: 029430-421

CURRENT APPLICATION NUMBER: US/09/699,684

CURRENT FILING DATE: 2000-10-31

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578

PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07

NUMBER OF SEQ ID NOS: 41

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5

LENGTH: 30

TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(30)
US-09-699-684-5

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||
Db 17 GVLAAL 22

RESULT 658

US-09-699-684-13
; Sequence 13, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(30)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-699-684-13

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||
Db 17 GVLAAL 22

RESULT 659

US-09-699-684-21
; Sequence 21, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 21
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(30)

; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-21

Query Match 5.1%; Score 6; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||
Db 17 GVLAAL 22

RESULT 660

US-09-348-578-6
; Sequence 6, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
; NAME/KEY: SIGNAL
; LOCATION: (1)..(31)
US-09-348-578-6

Query Match 5.1%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||
Db 18 GVLAAL 23

RESULT 661

US-09-348-578-14
; Sequence 14, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07

; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(31)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; US-09-348-578-14

Query Match 5.1%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 18 GVLAAL 23

RESULT 662

US-09-348-578-22
; Sequence 22, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(31)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; US-09-348-578-22

Query Match 5.1%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 18 GVLAAL 23

RESULT 663

US-09-699-684-6
; Sequence 6, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke

; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(31)
; US-09-699-684-6

Query Match 5.1%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 18 GVLAAL 23

RESULT 664

US-09-699-684-14
; Sequence 14, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(31)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; US-09-699-684-14

Query Match 5.1%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 18 GVLAAL 23

RESULT 665

US-09-699-684-22
; Sequence 22, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru

APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/699,684
CURRENT FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 31
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(31)
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
OTHER INFORMATION: signal
US-09-699-684-22

Query Match 5.1%; Score 6; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 18 GVLAAL 23

RESULT 666

US-09-348-578-7
Sequence 7, Application US/09348578
Patent No. 6160089
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
EARLIER APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 32
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
OTHER INFORMATION: signal
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(32)
US-09-348-578-7

Query Match 5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 19 GVLAAL 24

RESULT 667

US-09-348-578-15

Sequence 15, Application US/09348578
Patent No. 6160089
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
EARLIER APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15
LENGTH: 32
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(32)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
OTHER INFORMATION: signal
US-09-348-578-15

Query Match 5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 19 GVLAAL 24

RESULT 668

US-09-348-578-23
Sequence 23, Application US/09348578
Patent No. 6160089
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
EARLIER APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 32
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(32)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
OTHER INFORMATION: signal
US-09-348-578-23

Query Match 5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||

```
Db      19  GVLAAL 24

RESULT 669
US-09-699-684-7
; Sequence 7, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; NAME/KEY: SIGNAL
; LOCATION: (1)..(32)
US-09-699-684-7

Query Match      5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred.No.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19  GVLAAL 24
      19  GVLAAL 24
      |||||

RESULT 670
US-09-699-684-15
; Sequence 15, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(32)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-15

Query Match      5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred.No.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19  GVLAAL 24
      19  GVLAAL 24
      |||||

RESULT 671
US-09-699-684-23
; Sequence 23, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(32)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-23

Query Match      5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred.No.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19  GVLAAL 24
      19  GVLAAL 24
      |||||

RESULT 672
US-08-571-643A-7
; Sequence 7, Application US/08571643A
; Patent No. 5714371
; GENERAL INFORMATION:
; APPLICANT: Ramanathan, Lata
; APPLICANT: Wendel, Michelle
; TITLE OF INVENTION: Method for Refolding Insoluble
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/571,643A
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/439,680
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
```

```
Qy      19  GVLAAL 24
      19  GVLAAL 24
      |||||

Db      19  GVLAAL 24
      19  GVLAAL 24
      |||||

RESULT 671
US-09-699-684-23
; Sequence 23, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(32)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-23

Query Match      5.1%; Score 6; DB 2; Length 32;
Best Local Similarity 100.0%; Pred.No.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19  GVLAAL 24
      19  GVLAAL 24
      |||||

Db      19  GVLAAL 24
      19  GVLAAL 24
      |||||

RESULT 672
US-08-571-643A-7
; Sequence 7, Application US/08571643A
; Patent No. 5714371
; GENERAL INFORMATION:
; APPLICANT: Ramanathan, Lata
; APPLICANT: Wendel, Michelle
; TITLE OF INVENTION: Method for Refolding Insoluble
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/571,643A
; FILING DATE: 13-DEC-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/439,680
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
```

NAME: Dulak, No. 5714371man C.
REGISTRATION NUMBER: 31,308
REFERENCE/DOCKET NUMBER: JB0508K
TELEPHONE: 908-298-2906
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
FEATURE:
NAME/KEY: Carboxl 33 mer of NS4A
US-08-571-643A-7

Query Match 5.1%; Score 6; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CWWIVG 38
Db 1 CWWIVG 6

RESULT 673
US-08-571-643A-8
Sequence 8, Application US/08571643A
Patent No. 5714371
GENERAL INFORMATION:
APPLICANT: Ramanathan, Lata
APPLICANT: Wendel, Michelle
TITLE OF INVENTION: Method for Refolding Insoluble
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering-Plough Corporation
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.1
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/571,643A
FILING DATE: 13-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/439,680
FILING DATE: 12-MAY-1995
ATTORNEY/AGENT INFORMATION:
NAME: Dulak, No. 5714371man C.
REGISTRATION NUMBER: 31,308
REFERENCE/DOCKET NUMBER: JB0508K
TELEPHONE: 908-298-2906
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
FEATURE:
NAME/KEY: Carboxl 33 mer of NS4A of HCV-BK strain
US-08-571-643A-8

Query Match 5.1%; Score 6; DB 1; Length 33;

Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 PAIVPD 51
Db 14 PAIVPD 19
RESULT 674
US-08-439-747A-29
Sequence 29, Application US/08439747A
Patent No. 5767233
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Murray, Michael
APPLICANT: Ramanathan, Lata
TITLE OF INVENTION: Soluble, Cleavable Substrates of the Hepatitis
TITLE OF INVENTION: C Protease
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schering Corp.
STREET: 2000 Galloping Hill Road
CITY: Kenilworth
STATE: New Jersey
COUNTRY: USA
ZIP: 07033-0530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5.3
SOFTWARE: Microsoft Word 5.1a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/439,747A
FILING DATE: May 12, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lunn, Paul G.
REGISTRATION NUMBER: 32,743
REFERENCE/DOCKET NUMBER: JB0509
TELEPHONE: 908-298-5061
TELEFAX: 908-298-5388
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acid residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
FEATURE:
NAME/KEY: Carboxl 33 mer of NS4A
US-08-439-747A-29

Query Match 5.1%; Score 6; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CWWIVG 38
Db 1 CWWIVG 6

RESULT 675
US-08-439-747A-30
Sequence 30, Application US/08439747A
Patent No. 5767233
GENERAL INFORMATION:
APPLICANT: Zhang, Rumin
APPLICANT: Murray, Michael
APPLICANT: Ramanathan, Lata
TITLE OF INVENTION: Soluble, Cleavable Substrates of the Hepatitis
TITLE OF INVENTION: C Protease
NUMBER OF SEQUENCES: 34

```
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering Corp.
;; STREET: 2000 Galloping Hill Road
;; CITY: Kenilworth
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07033-0530
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: Apple Macintosh
;; OPERATING SYSTEM: Macintosh 7.5.3
;; SOFTWARE: Microsoft Word 5.1a
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/439,747A
;; FILING DATE: May 12, 1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lunn, Paul G.
;; REGISTRATION NUMBER: 32,743
;; REFERENCE/DOCKET NUMBER: JB0509
;; TELEPHONE: 908-298-5061
;; TELEFAX: 908-298-5388
;; INFORMATION FOR SEQ ID NO: 30:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 33 amino acid residues
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: polypeptide
;; FEATURE:
;; NAME/KEY: Carboxyl 33 mer of NS4A of HCV-BK strain
US-08-439-747A-30

Query Match 5.1%; Score 6; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 PAIPVD 51
Db 14 PAIPVD 19

RESULT 676
US-08-440-409B-30
; Sequence 30, Application US/08440409B
; Patent No. 5843752
; GENERAL INFORMATION:
; APPLICANT: Dasmahapatra, Bimal
; APPLICANT: Murray, Michael
; APPLICANT: Ramanathan, Lata
; APPLICANT: Butkiewicz, Nancy
; TITLE OF INVENTION: Soluble Active Hepatitis C Virus Protease
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5.3
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,409B
; FILING DATE: May 12, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: JB0494
; TELEPHONE: 908-298-5061
; TELEFAX: 908-298-5388
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Carboxyl 33 mer of NS4A of HCV-BK strain
US-08-440-409B-31

Query Match 5.1%; Score 6; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CWVIVG 38
Db 1 CWVIVG 6

RESULT 677
US-08-440-409B-31
; Sequence 31, Application US/08440409B
; Patent No. 5843752
; GENERAL INFORMATION:
; APPLICANT: Dasmahapatra, Bimal
; APPLICANT: Murray, Michael
; APPLICANT: Ramanathan, Lata
; APPLICANT: Butkiewicz, Nancy
; TITLE OF INVENTION: Soluble Active Hepatitis C Virus Protease
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5.3
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,409B
; FILING DATE: May 12, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: JB0494
; TELEPHONE: 908-298-5061
; TELEFAX: 908-298-5388
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: polypeptide
; FEATURE:
; NAME/KEY: Carboxyl 33 mer of NS4A of HCV-BK strain
US-08-440-409B-31

Query Match 5.1%; Score 6; DB 1; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 46 PAIVPD 51
Db 14 PAIVPD 19

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 678
US-09-348-578-8
; Sequence 8, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(33)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-8

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 20 GVLAAL 25

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 679
US-09-348-578-16
; Sequence 16, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(33)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-16

QY 19 GVLAAL 24
Db 20 GVLAAL 25

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 680
US-09-348-578-24
; Sequence 24, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(33)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-24

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 20 GVLAAL 25

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 681
US-09-699-684-8
; Sequence 8, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL

; LOCATION: (1)...(33)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-8

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 20 GVLAAL 25

RESULT 682

US-09-699-684-16
; Sequence 16, Application US/09699684
; Patent No. 6436674

GENERAL INFORMATION:

; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 33

ORGANISM: Artificial Sequence

FEATURE:

NAME/KEY: SIGNAL

LOCATION: (1)...(33)

; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-16

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 20 GVLAAL 25

RESULT 683

US-09-699-684-24
; Sequence 24, Application US/09699684
; Patent No. 6436674

GENERAL INFORMATION:

; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 33

TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)...(33)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-24

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 20 GVLAAL 25

RESULT 684

US-08-700-356-3
; Sequence 3, Application US/08700356
; Patent No. 5739002

GENERAL INFORMATION:

; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: FAILLA, Cristina
; APPLICANT: TOMEI, Licia
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE
; TITLE OF INVENTION: PROTEOLYTIC ACTIVITY OF THE NS3 HEPATITIS C VIRUS (HCV)
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/700,356
; FILING DATE: 23-AUG-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: DE FRANCESCO=1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197

TELEFAX: 202-737-3528

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

; LENGTH: 34 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

FEATURE:

; OTHER INFORMATION: /note= "Xaa at position 34 means

; OTHER INFORMATION: Abu (2-Aminobutyric acid)"

US-08-700-356-3

Query Match 5.1%; Score 6; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 PAIVPD 51
| | | | |
Db 15 PAIVPD 20

RESULT 685


```
US-08-936-865-3
; Sequence 3, Application US/08936865
; Patent No. 5861297
; GENERAL INFORMATION:
; APPLICANT: Sardana, Vinod V
; APPLICANT: Blue, Jeffrey T
; TITLE OF INVENTION: DETERGENT-FREE HEPATITIS C PROTEASE
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.
; CITY: Rahway
; STATE: NJ
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,865
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ayler, Sylvia A
; REGISTRATION NUMBER: 36,436
; REFERENCE/DOCKET NUMBER: 19691
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-4909
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLSCULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; LIBRARY: Cofactor of NS3 serine protease
; CLONE: Solid phase peptide synthesis
; US-08-936-865-3

Query Match 5.1%; Score 6; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 PAIVPD 51
Db 15 PAIVPD 20

RESULT 686
US-09-348-578-9
; Sequence 9, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; EARLIER FILING DATE: 1999-07-07
; CURRENT FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; US-09-348-578-9

Query Match 5.1%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 21 GVLAAL 26

RESULT 687
US-09-348-578-17
; Sequence 17, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; EARLIER FILING DATE: 1999-07-07
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; US-09-348-578-17

Query Match 5.1%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 21 GVLAAL 26

RESULT 688
US-09-348-578-25
; Sequence 25, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
```

```
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(34)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-25

Query Match          5.1%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GVLAAL 24
        |||||
Db      21 GVLAAL 26

RESULT 689
US-09-699-684-9
; Sequence 9, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(34)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-9

Query Match          5.1%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GVLAAL 24
        |||||
Db      21 GVLAAL 26

RESULT 689
US-09-699-684-17
; Sequence 17, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/348,578
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(34)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-17

Query Match          5.1%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GVLAAL 24
        |||||
Db      21 GVLAAL 26

RESULT 690
US-09-699-684-25
; Sequence 25, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER APPLICATION NUMBER: 09/348,578
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(34)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-25

Query Match          5.1%; Score 6; DB 2; Length 34;
Best Local Similarity 100.0%; Pred.No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      19 GVLAAL 24
        |||||
Db      21 GVLAAL 26

RESULT 692
US-09-348-578-18
; Sequence 18, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
```

APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
EARLIER APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(35)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-348-578-18

Query Match 5.1%; Score 6; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 22 GVLAAL 27

RESULT 693

US-09-348-578-26
Sequence 26, Application US/09348578
Patent No. 6160089
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/348,578
CURRENT FILING DATE: 1999-07-07
EARLIER APPLICATION NUMBER: JP 193003/1998
EARLIER FILING DATE: 1998-07-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(35)
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-348-578-26

Query Match 5.1%; Score 6; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 22 GVLAAL 27

RESULT 694

US-09-699-684-18
Sequence 18, Application US/09699684
Patent No. 6436674
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/699,684
CURRENT FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(35)
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-699-684-18

Query Match 5.1%; Score 6; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 22 GVLAAL 27

RESULT 695

US-09-699-684-26
Sequence 26, Application US/09699684
Patent No. 6436674
GENERAL INFORMATION:
APPLICANT: HONJO, Masaru
APPLICANT: NAITOH, Naokazu
APPLICANT: UCHIDA, Hiroshi
APPLICANT: MOCHIZUKI, Daisuke
APPLICANT: MATSUMOTO, Kazuya
TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
FILE REFERENCE: 029430-421
CURRENT APPLICATION NUMBER: US/09/699,684
CURRENT FILING DATE: 2000-10-31
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 26
LENGTH: 35
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(35)
OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
US-09-699-684-26

Query Match 5.1%; Score 6; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 22 GVLAAL 27

RESULT 696

US-09-348-578-27
; Sequence 27, Application US/09348578
; Patent No. 6160089
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/348,578
; CURRENT FILING DATE: 1999-07-07
; EARLIER APPLICATION NUMBER: JP 193003/1998
; EARLIER FILING DATE: 1998-07-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(36)
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-348-578-27

Query Match 5.1%; Score 6; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 23 GVLAAL 28

RESULT 697

US-09-699-684-27
; Sequence 27, Application US/09699684
; Patent No. 6436674
; GENERAL INFORMATION:
; APPLICANT: HONJO, Masaru
; APPLICANT: NAITOH, Naokazu
; APPLICANT: UCHIDA, Hiroshi
; APPLICANT: MOCHIZUKI, Daisuke
; APPLICANT: MATSUMOTO, Kazuya
; TITLE OF INVENTION: METHOD FOR SECRETORY PRODUCTION OF HUMAN GROWTH HORMONE
; FILE REFERENCE: 029430-421
; CURRENT APPLICATION NUMBER: US/09/699,684
; CURRENT FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/348,578
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(36)
; OTHER INFORMATION: Description of Artificial Sequence:Modified OppA secretion
; OTHER INFORMATION: signal
US-09-699-684-27

Query Match 5.1%; Score 6; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 23 GVLAAL 28

RESULT 698

US-09-595-682B-2
; Sequence 2, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/031171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-595-682B-2

Query Match 5.1%; Score 6; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
Db 12 KGVVLG 17

RESULT 699

US-09-902-540-16625
; Sequence 16625, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16625
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16625

Query Match 5.1%; Score 6; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGLLQ 91
Db 30 VLGLLQ 35

RESULT 700

US-08-118-270-185

Sequence 185, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 185:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-185

Query Match 5.1%; Score 6; DB 1; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 5 VLLGGV 10
|||||

RESULT 701
PCT-US93-08528-185
Sequence 185, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

Sequence 185, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
APPLICANT: Schuster, David I.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 185:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-185

Query Match 5.1%; Score 6; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 5 VLLGGV 10
|||||

RESULT 702
US-08-617-929-9
Sequence 9, Application US/08617929
Patent No. 585771
GENERAL INFORMATION:
APPLICANT: KUMAZAWA, Toshiaki
TITLE OF INVENTION: ANTIGENIC PEPTIDE COMPOUND AND
TITLE OF INVENTION: IMMUNOASSAY
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/617,929
FILING DATE: 24-APR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/JP94/01823
FILING DATE: 28-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 6/207695
FILING DATE: 31-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5/272864
FILING DATE: 29-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 77384/109
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-617-929-9

Query Match 5.1%; Score 6; DB 1; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 PAIVPD 51
Db 7 PAIVPD 12

RESULT 703
US-08-537-811-44
; Sequence 44, Application US/08537811
; Patent No. 5910405
; GENERAL INFORMATION:
; APPLICANT: CHO, JOONG MYUNG
; APPLICANT: CHOI, DEOG YOUNG
; APPLICANT: KIM, CHUN HYUNG
; APPLICANT: SO, HONG SEOB
; APPLICANT: YANG, JAE YOUNG
; APPLICANT: KIM, IN SOO
; APPLICANT: KIM, JOO HO
; TITLE OF INVENTION: IMPROVED HCV DIAGNOSTIC
; TITLE OF INVENTION: AGENTS
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,811
; FILING DATE: 24-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/KR94/00040
; FILING DATE: 29-APR-1994
; APPLICATION NUMBER: KR 93-7440
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jones, Ili Harry C
; REGISTRATION NUMBER: 20,280
; REFERENCE/DOCKET NUMBER: 8512-037-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; OTHER INFORMATION: KHCV NS4B, Fig. 2
US-08-537-811-44

Query Match 5.1%; Score 6; DB 1; Length 42;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 60 DEMEEC 65
Db 13 DEMEEC 18

RESULT 704
US-07-946-054-10
; Sequence 10, Application US/07946054
; Patent No. 5582968
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; APPLICANT: Hosein, Barbara H
; TITLE OF INVENTION: NO. 5582968a1 Branched Hybrid and Cluster
; TITLE OF INVENTION: Peptides Effective in Diagnosing and Detecting No. 5582968-B Hepatitis
; TITLE OF INVENTION: No. 5582968-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: United Biomedical Inc.
; STREET: 25 Davids Dr.
; CITY: Hauppauge
; STATE: New York
; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/946,054
; FILING DATE: 15-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, M. Lisa
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 2000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-273-2828
; TELEFAX: 516-273-1717
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-946-054-10

Query Match 5.1%; Score 6; DB 1; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DEMEEC 65
Db 18 DEMEEC 23

RESULT 705
PCT-US93-08638-10
; Sequence 10, Application PC/TUS9308638
; GENERAL INFORMATION:
; APPLICANT: United Biomedical Inc.
; TITLE OF INVENTION: Novel Branched Hybrid and Cluster Peptides
; TITLE OF INVENTION: Effective in Diagnosing and Detecting Non-A,
; TITLE OF INVENTION: Non-B Hepatitis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: UNITED BIOMEDICAL INC.
; STREET: 25 Davids Drive
; CITY: Hauppauge
; STATE: New York

```

; COUNTRY: USA
; ZIP: 11788
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08638
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: M. Lisa Wilson
; REGISTRATION NUMBER: 34,045
; REFERENCE/DOCKET NUMBER: 9055
; TELEPHONE: 516-273-2828
; TELEFAX: 516-273-1717
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 47 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08638-10

Query Match 5.1%; Score 6; DB 4; Length 47;
Best Local Similarity 100.0%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 60 DEMEEC 65
DB 18 DEMEEC 23

RESULT 706
US-09-595-682B-4
; Sequence 4, Application US/09595682B
; Patent No. 6800483
; GENERAL INFORMATION:
; APPLICANT: Danks, Mary K.
; APPLICANT: Potter, Philip M.
; APPLICANT: Houghton, Peter J.
; TITLE OF INVENTION: Compositions and Methods for Sensitizing and Inhibiting Growth of
; TITLE OF INVENTION: Tumor Cells
; FILE REFERENCE: SJ-0005
; CURRENT APPLICATION NUMBER: US/09/595,682B
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/075,258
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: PCT/US99/03171
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-09-595-682B-4

Query Match 5.1%; Score 6; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 83 KGVILG 88
DB 30 KGVILG 35

RESULT 707
US-08-905-223-407
; Sequence 407, Application US/08905223
```

```

; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duelt, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 407:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 57 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -47...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 7.9
; OTHER INFORMATION: seq LLLPRVLLTMASG/SP
; US-08-905-223-407

Query Match 5.1%; Score 6; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 15 VLLGGV 20
DB 9 VLLGGV 14

RESULT 708
US-08-685-764-1
; Sequence 1, Application US/08685764
; Patent No. 5800982
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, AKIRA
; APPLICANT: MAKI, NORU
; APPLICANT: YAGI, SHINTARO
; APPLICANT: KASHIWAKURA, TOMIKO
; TITLE OF INVENTION: ANTIGENIC PEPTIDES FOR GROUPING
; TITLE OF INVENTION: HEPATITIS C VIRUS, KIT COMPRISING THE SAME AND
; TITLE OF INVENTION: METHODS FOR ITS GROUPING USING THE SAME
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
```

;; CITY: BOSTON
;; STATE: MASSACHUSETTS
;; COUNTRY: US
;; ZIP: 02109
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/685,764
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/092,192
;; FILING DATE: 15-JUL-1993
;; APPLICATION NUMBER: JP 212061/92
;; FILING DATE: 16-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 316634/92
;; FILING DATE: 30-OCT-1992
;; APPLICATION NUMBER: JP 316635/92
;; FILING DATE: 30-OCT-1992
;; APPLICATION NUMBER: JP 104754/93
;; FILING DATE: 30-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BUCKLEY, LINDA M.
;; REGISTRATION NUMBER: 31003
;; REFERENCE/DOCKET NUMBER: 42822
;; TELEPHONE: (617) 523-3400
;; TELEFAX: (617) 523-6440
;; TELEX: 200291 STRE UR
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 63 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-685-764-1

Query Match 5.1%; Score 6; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
DB 33 DEMEEC 38

RESULT 709
US-08-685-764-3
; Sequence 3, Application US/08685764
; Patent No. 5800982
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, AKIRA
; APPLICANT: MAKI, NORIURU
; APPLICANT: YAGI, SHINTARO
; APPLICANT: KASHIWAKUMA, TOMIKO
; TITLE OF INVENTION: ANTI-GEN PEPTIDES FOR GROUPING
; TITLE OF INVENTION: HEPATITIS C VIRUS, KIT COMPRISING THE SAME AND
; TITLE OF INVENTION: METHODS FOR ITS GROUPING USING THE SAME
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/685,764
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/092,192
;; FILING DATE: 15-JUL-1993
;; APPLICATION NUMBER: JP 212061/92
;; FILING DATE: 16-JUL-1992
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: JP 316634/92
;; FILING DATE: 30-OCT-1992
;; APPLICATION NUMBER: JP 316635/92
;; FILING DATE: 30-OCT-1992
;; APPLICATION NUMBER: JP 104754/93
;; FILING DATE: 30-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BUCKLEY, LINDA M.
;; REGISTRATION NUMBER: 31003
;; REFERENCE/DOCKET NUMBER: 42822
;; TELEPHONE: (617) 523-3400
;; TELEFAX: (617) 523-6440
;; TELEX: 200291 STRE UR
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 63 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-685-764-3

Query Match 5.1%; Score 6; DB 1; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEEC 65
DB 33 DEMEEC 38

RESULT 710
US-09-513-999C-6688
; Sequence 6688, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59 US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 6688
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-6688

Query Match 5.1%; Score 6; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;


```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 109 OKLEAF 114
Db 52 OKLEAF 57

RESULT 711
US-09-248-796A-25370
; Sequence 25370, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25370
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-25370

Query Match 5.1%; Score 6; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 86 VLGLLQ 91
Db 56 VLGLLQ 61

RESULT 712
US-09-513-999C-7755
; Sequence 7755, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7755
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-7755

Query Match 5.1%; Score 6; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 109 OKLEAF 114
Db 52 OKLEAF 57

RESULT 713
US-09-902-540-13136
; Sequence 13136, Application US/09902540
```

```
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13136
; LENGTH: 78
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13136

Query Match 5.1%; Score 6; DB 2; Length 78;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALA 25
Db 14 VLAALA 19

RESULT 714
US-08-685-764-2
; Sequence 2, Application US/08685764
; Patent No. 5800982
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, AKIRA
; APPLICANT: MAKI, NOBORU
; APPLICANT: YAGI, SHINTARO
; APPLICANT: KASHIWAKUMA, TOMIKO
; TITLE OF INVENTION: ANTIGENIC PEPTIDES FOR GROUPING
; TITLE OF INVENTION: HEPATITIS C VIRUS, KIT COMPRISING THE SAME AND
; TITLE OF INVENTION: METHODS FOR ITS GROUPING USING THE SAME
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSER: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSER: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,764
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/092,192
; FILING DATE: 15-JUL-1993
; APPLICATION NUMBER: JP 212061/92
; FILING DATE: 16-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316634/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 316635/92
; FILING DATE: 30-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 104754/93
; FILING DATE: 30-APR-1993
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: BUCKLEY, LINDA M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 42822
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 87 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-685-764-2

Query Match          5.1%; Score 6; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEEC 65
Db      33 DEMEEC 38

RESULT 715
US-08-479-078-1
; Sequence 1, Application US/08479078
; Patent No. 5814466
; GENERAL INFORMATION:
; APPLICANT: Pawson, Anthony
; TITLE OF INVENTION: Method for Assaying for a Substance that
; TITLE OF INVENTION: Affects an SH2-Phosphorylated Ligand Regulatory System
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street, West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,078
; FILING DATE: June 6, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Linda M. Kurdzyk
; REGISTRATION NUMBER: 34,971
; REFERENCE/DOCKET NUMBER: 3153-154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-479-078-1

Query Match          5.1%; Score 6; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      26 AYCLSV 31
Db      36 AYCLSV 41

NAME: BUCKLEY, LINDA M.
REGISTRATION NUMBER: 31003
REFERENCE/DOCKET NUMBER: 42822
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 87 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-685-764-2

Query Match          5.1%; Score 6; DB 1; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      60 DEMEEC 65
Db      33 DEMEEC 38

RESULT 716
US-09-107-532A-3655
; Sequence 3655, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 3655:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...97
; SEQUENCE DESCRIPTION: SEQ ID NO: 3655:
US-09-107-532A-3655

Query Match          5.1%; Score 6; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 LAALAA 26
Db      66 LAALAA 71

RESULT 717
US-08-308-086-3
; Sequence 3, Application US/08308086
; Patent No. 5786454
; GENERAL INFORMATION:
; APPLICANT: Gabriel Waksman and Andrey Shaw
; TITLE OF INVENTION: Modified SH2 Domains and Methods of Use
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
```

; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/308,086
; FILING DATE: 09-16-97
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Padst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU105
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-308-086-3

Query Match 5.1%; Score 6; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 36 AYCLSV 41

RESULT 718
US-08-975-040-21
; Sequence 21, Application US/08975040
; Patent No. 6251620
; GENERAL INFORMATION:
; APPLICANT: HATADA, MARCOS
; APPLICANT: LU, XIAODE
; APPLICANT: LAIRD, ELLEN
; APPLICANT: KARAS, JENNIFER
; APPLICANT: ZOLLER, MARK
; APPLICANT: HOLT, DENNIS
; TITLE OF INVENTION: MACHINE READABLE STORAGE MEDIUM RELATING
; TO ZAP-FAMILY PROTEINS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID L. BERSTEIN, ARIAD PHARMACEUTICALS,
; ADDRESSEE: INC.
; STREET: 26 LANDSDOWNE STREET
; CITY: CAMBRIDGE
; STATE: MA
; COUNTRY: US
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,040
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,578
; FILING DATE: 22-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BERSTEIN, DAVID L.

; REGISTRATION NUMBER: 31,235
; REFERENCE/DOCKET NUMBER: ARIAD 347P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-494-0400
; TELEFAX: 617-494-1828
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-975-040-21

Query Match 5.1%; Score 6; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 36 AYCLSV 41

RESULT 719
US-08-167-035-19
; Sequence 19, Application US/08167035
; Patent No. 5618691
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; KINASES AND NOVEL TARGET PROTEINS
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: 10036-2711
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-167-035-19

Query Match 5.1%; Score 6; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31


```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-539-005-19

Query Match 5.1%; Score 6; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 36 AYCLSV 41

RESULT 723
US-09-280-598-21
; Sequence 21, Application US/09280598
; Patent No. 6391584
; GENERAL INFORMATION:
; APPLICANT: Schlusser, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; APPLICANT: App. Harold
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; KINASES AND NOVEL TARGET PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/280,598
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/252,820
; FILING DATE: 02-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-067
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-280-598-21

Query Match 5.1%; Score 6; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 36 AYCLSV 41

US-09-087-465-37
; Sequence 37, Application US/09087465A
; Patent No. 6160092
; GENERAL INFORMATION:
; APPLICANT: Vinkemeier, Uwe
; APPLICANT: Chen, Xiaomin
; APPLICANT: Darnell Jr., James E
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF THE CORE PORTION OF A STAT AND METHODS OF
; TITLE OF INVENTION: USE
; FILE REFERENCE: 600-1-229
; CURRENT APPLICATION NUMBER: US/09/087,465A
; CURRENT FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 37
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Rous sarcoma virus
US-09-087-465-37

Query Match 5.1%; Score 6; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 37 AYCLSV 42

RESULT 725
US-09-513-999C-6858
; Sequence 6858, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent.pm
; SEQ ID NO 6858
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 70
; OTHER INFORMATION: Xaa=Gly or Val
US-09-513-999C-6858

Query Match 5.1%; Score 6; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LIQRAT 94
Db 48 LIQRAT 53

RESULT 726
```

US-08-820-754-22
; Sequence 22, Application US/08820754
; Patent No. 5976835
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/820,754
; FILING DATE: 19-MAR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,185
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; CLONE: Src
; PUBLICATION INFORMATION:
; AUTHORS: Waksman, et al.
; JOURNAL: Nature
; VOLUME: 358
; PAGES: 646-653
; DATE: 1992
US-08-820-754-22

Query Match 5.1%; Score 6; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 26 AYCLSV 31
Db 39 AYCLSV 44
RESULT 727
US-08-956-652-22
; Sequence 22, Application US/08956652
; Patent No. 6013475
; GENERAL INFORMATION:
; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,652
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,185
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:
; CLONE: Src
; PUBLICATION INFORMATION:
; AUTHORS: Waksman, et al.
; JOURNAL: Nature
; VOLUME: 358

;
; PAGES: 646-653
; DATE: 1992
US-08-956-652-22

Query Match 5.1%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 39 AYCLSV 44

RESULT 728

US-08-956-869-22
; Sequence 22, Application US/08956869
; Patent No. 6030808

GENERAL INFORMATION:

; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/956,869
; FILING DATE:

CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/212,185
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993

ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521

INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; IMMEDIATE SOURCE:

;
; CLONE: Src
; PUBLICATION INFORMATION:
; AUTHORS: Waksman, et al.
; JOURNAL: Nature
; VOLUME: 358
; PAGES: 646-653
; DATE: 1992
US-08-956-869-22

Query Match 5.1%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 39 AYCLSV 44

RESULT 729

US-08-948-547-22
; Sequence 22, Application US/08948547
; Patent No. 6124118

GENERAL INFORMATION:

; APPLICANT: Darnell Jr., James E.
; APPLICANT: Schindler, Christian W.
; APPLICANT: Fu, Xian-Yuan
; APPLICANT: Wen, Zilong
; APPLICANT: Zhong, Zhong
; TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
; TITLE OF INVENTION: SEQUENCES AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/948,547
; FILING DATE:

CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,185
; FILING DATE: 11-MAR-1994
; APPLICATION NUMBER: US 07/980,498
; FILING DATE: 23-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/854,296
; FILING DATE: 19-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO US93/02569
; FILING DATE: 19-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,588
; FILING DATE: 24-SEP-1993

ATTORNEY/AGENT INFORMATION:

; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-073 CIP
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 amino acids
; TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
IMMEDIATE SOURCE:
CLONE: Src
PUBLICATION INFORMATION:
AUTHORS: Wakeman, et al.
JOURNAL: Nature
VOLUME: 358
PAGES: 646-653
DATE: 1992
US-08-948-547-22

Query Match 5.1%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
Db 39 AYCLSV 44

RESULT 730

US-09-732-210-371
Sequence 371, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mittanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 371
LENGTH: 105
TYPE: PRT
ORGANISM: Candida maltosa
US-09-732-210-371

Query Match 5.1%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELGG 44
Db 89 HIELGG 94

RESULT 731

US-09-732-210-1056
Sequence 1056, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mittanck, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07

PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 1056
LENGTH: 105
TYPE: PRT
ORGANISM: Candida tropicalis
US-09-732-210-1056

Query Match 5.1%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELGG 44
Db 89 HIELGG 94

RESULT 732

US-08-212-185-22
Sequence 22, Application US/08212185
Patent No. 6605442
GENERAL INFORMATION:
APPLICANT: Darnell Jr., James E.
APPLICANT: Schindler, Christian W.
APPLICANT: Fu, Xian-Yuan
APPLICANT: Wen, Zilong
APPLICANT: Zhong, Zhong
TITLE OF INVENTION: RECEPTOR RECOGNITION FACTORS, PROTEIN
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,185
FILING DATE: 11-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/980,498
FILING DATE: 23-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/854,296
FILING DATE: 19-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO US93/02569
FILING DATE: 19-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/126,588
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-073 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 amino acids
TYPE: amino acid

;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; HYPOTHETICAL: NO
;; ANTI-SENSE: NO
;; FRAGMENT TYPE: internal
;; IMMEDIATE SOURCE:
;; CLONE: Src
;; PUBLICATION INFORMATION:
;; AUTHORS: Waksman, et al.
;; JOURNAL: Nature
;; VOLUME: 358
;; PAGES: 646-653
;; DATE: 1992
US-08-212-185-22

Query Match 5.1%; Score 6; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
Db 39 AYCLSV 44

RESULT 733
US-08-858-207A-273
; Sequence 273, Application US/08858207A
; Patent No. 6348328
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Nicholas, Richard
; APPLICANT: Stodola, Robert
; TITLE OF INVENTION: No. 6348328el Compounds
; NUMBER OF SEQUENCES: 552
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/858,207A
; FILING DATE: 09-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/017670
; FILING DATE: 14-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50475
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 273:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6348328e
US-08-858-207A-273

Query Match 5.1%; Score 6; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLEVTT 11
Db 10 DLEVTT 15

RESULT 734
US-09-417-251A-2
; Sequence 2, Application US/09417251A
; Patent No. 6720474
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Herrman, Rafael
; APPLICANT: Rafalski, Antoni
; APPLICANT: McCutchen, Bill F.
; TITLE OF INVENTION: Plant Protein Disulfide Isomerases
; FILE REFERENCE: BB1085 US NA
; CURRENT APPLICATION NUMBER: US/09/417,251A
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/049,408
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Zea mays
US-09-417-251A-2

Query Match 5.1%; Score 6; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAY 27
Db 2 AALAAAY 7

RESULT 735
US-09-417-251A-2
; Sequence 2, Application US/09417251A
; Patent No. 6864403
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Herrman, Rafael
; APPLICANT: Rafalski, Antoni
; APPLICANT: McCutchen, Bill F.
; TITLE OF INVENTION: Plant Protein Disulfide Isomerases
; FILE REFERENCE: BB1085 US NA
; CURRENT APPLICATION NUMBER: US/09/417,251A
; CURRENT FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/049,408
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Zea mays
US-09-417-251A-2

Query Match 5.1%; Score 6; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAY 27
Db 2 AALAAAY 7

RESULT 736

US-09-504-358-24
; Sequence 24, Application US/09504358
; Patent No. 6365376
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E.
; APPLICANT: Brzostowicz, Patricia C.
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATES
; FILE REFERENCE: BC1001 US NA
; CURRENT APPLICATION NUMBER: US/09/504,358
; CURRENT FILING DATE: 2000-02-15
; EARLIER APPLICATION NUMBER: 60/120,702
; EARLIER FILING DATE: 1999-February-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Brevibacterium sp HCU
US-09-504-358-24

Query Match 5.1%; Score 6; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0

Qy 20 VLAALA 25
| | | | |
Db 17 VLAALA 22

RESULT 737

US-09-954-314-24
; Sequence 24, Application US/09954314
; Patent No. 6465324
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E.
; APPLICANT: Brzostowicz, Patricia C.
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATES
; FILE REFERENCE: BC1001 US NA
; CURRENT APPLICATION NUMBER: US/09/954,314
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-February-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Brevibacterium sp HCU
US-09-954-314-24

Query Match 5.1%; Score 6; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0

Qy 20 VLAALA 25
| | | | |
Db 17 VLAALA 22

RESULT 738

US-10-230-562-24
; Sequence 24, Application US/10230562
; Patent No. 6790645
; GENERAL INFORMATION:
; APPLICANT: Rouviere, Pierre E.
; APPLICANT: Brzostowicz, Patricia C.
; TITLE OF INVENTION: GENES AND ENZYMES FOR THE PRODUCTION OF ADIPIC ACID INTERMEDIATES
; FILE REFERENCE: BC-1001
; CURRENT APPLICATION NUMBER: US/10/230,562

; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/120,702
; PRIOR FILING DATE: 1999-02-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Brevibacterium sp HCU
US-10-230-562-24

Query Match 5.1%; Score 6; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 3.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0

Qy 20 VLAALA 25
| | | | |
Db 17 VLAALA 22

RESULT 739

US-09-583-110-3398
; Sequence 3398, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3398
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3398

Query Match 5.1%; Score 6; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVTT 11
| | | | |
Db 10 DLEVTT 15

RESULT 740

US-09-107-433-3440
; Sequence 3440, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>

;; OPERATING SYSTEM: <Unknown>
;; SOFTWARE: <Unknown>
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/107,433
;; FILING DATE: 30-Jun-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/ 085131
;; FILING DATE: May 12, 1998
;; APPLICATION NUMBER: 60/051553
;; FILING DATE: July 2, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ariniello, Pamela Deneke
;; REGISTRATION NUMBER: 40,489
;; REFERENCE/DOCKET NUMBER: GTC-011
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781)893-5007
;; TELEFAX: (781)893-8277
;; INFORMATION FOR SEQ ID NO: 3440:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 120 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Streptococcus pneumoniae
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (B) LOCATION 1...120
;; SEQUENCE DESCRIPTION: SEQ ID NO: 3440:
US-09-107-433-3440

Query Match 5.1%; Score 6; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLEVT 11
Db 15 DLEVT 20

RESULT 741
US-09-248-796A-17484
;; Sequence 17484, Application US/09248796A
;; Patent No. 6747137
;; GENERAL INFORMATION:
;; APPLICANT: Keith Weinstock et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
;; FILE REFERENCE: 107196.132
;; CURRENT APPLICATION NUMBER: US/09/248,796A
;; PRIOR FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074,725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096,409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 17484
;; LENGTH: 123
;; TYPE: PRT
;; ORGANISM: Candida albicans
US-09-248-796A-17484

Query Match 5.1%; Score 6; DB 2; Length 123;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 AVIEPI 103
Db 63 AVIEPI 68

RESULT 742

US-08-811-481-5
;; Sequence 5, Application US/08811481
;; Patent No. 6300093
;; GENERAL INFORMATION:
;; APPLICANT: Kinsdovogel, Wayne
;; APPLICANT: Jelinek, Laura J.
;; APPLICANT: Sheppard, Paul O.
;; APPLICANT: Hagopian, William A.
;; APPLICANT: LaGasse, James M.
;; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
;; NUMBER OF SEQUENCES: 34
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Fast-SEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/811,481
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lingenfelter, Susan
;; REGISTRATION NUMBER: P-41,156
;; REFERENCE/DOCKET NUMBER: 95-36
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6675
;; TELEFAX: 206-442-6678
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 127 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-811-481-5

Query Match 5.1%; Score 6; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWIV 37
Db 36 GCWIV 41

RESULT 743
US-09-876-527-5
;; Sequence 5, Application US/09876527
;; Patent No. 6627735
;; GENERAL INFORMATION:
;; APPLICANT: Kinsdovogel, Wayne
;; APPLICANT: Jelinek, Laura J.
;; APPLICANT: Sheppard, Paul O.
;; APPLICANT: Hagopian, William A.
;; APPLICANT: LaGasse, James M.
;; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
;; NUMBER OF SEQUENCES: 34
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA

```
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,527
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/811,481
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-876-527-5
Query Match 5.1%; Score 6; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 32 GCVVIV 37
Db 36 GCVVIV 41
RESULT 744
US-09-902-540-14345
; Sequence 14345, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14345
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14345
Query Match 5.1%; Score 6; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 LAALAA 26
Db 69 LAALAA 74
RESULT 745
US-09-134-000C-3900
; Sequence 3900, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3900
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-3900
Query Match 5.1%; Score 6; DB 2; Length 129;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALA 25
Db 82 VLAALA 87
RESULT 746
US-09-308-386A-4
; Sequence 4, Application US/09308386A
; Patent No. 6605704
; GENERAL INFORMATION:
; APPLICANT: Tatsuo, SUGIYAMA et al.
; TITLE OF INVENTION: PLANT REGULATOR PROTEIN AND NUCLEIC ACID CODING FOR THE SAME
; FILE REFERENCE: 0760-0266P
; CURRENT APPLICATION NUMBER: US/09/308,386A
; CURRENT FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 134
; TYPE: PRT
; ORGANISM: S.typhimurium
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(134)
; OTHER INFORMATION: Any Xaa = any amino acid, unknown or other
US-09-308-386A-4
Query Match 5.1%; Score 6; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALA 25
Db 40 VLAALA 45
RESULT 747
US-09-902-540-14756
; Sequence 14756, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
```

; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14756
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14756

Query Match 5.1%; Score 6; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||
DB 40 LLGGVL 45

RESULT 748
US-09-328-352-4915
; Sequence 4915, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4915
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4915

Query Match 5.1%; Score 6; DB 2; Length 143;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 IEQAQV 77
|||
DB 67 IEQAQV 72

RESULT 749
US-09-328-352-7270
; Sequence 7270, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7270
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7270

Query Match 5.1%; Score 6; DB 2; Length 144;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||
DB 70 LLGGVL 75

RESULT 750

US-09-902-540-12071
; Sequence 12071, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12071
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12071

Query Match 5.1%; Score 6; DB 2; Length 146;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLIQ 91
|||
DB 53 VLGLIQ 58

RESULT 751
US-09-252-991A-32059
; Sequence 32059, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32059
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32059

Query Match 5.1%; Score 6; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
DB 75 LAALAA 80

RESULT 752
US-09-270-767-49120
; Sequence 49120, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 49120
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-49120

Query Match          5.1%; Score 6; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      40 IELGK 45
Db      75 IELGK 80

RESULT 753
US-08-336-553A-2
; Sequence 2, Application US/08336553A
; Patent No. 6054264
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; ADDRESSEE: KUO, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,157
; FILING DATE: 11-May-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553A
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 155 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-439-157-2

Query Match          5.1%; Score 6; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      11 TSTWVL 16
Db      65 TSTWVL 70

RESULT 755
US-09-437-895-2
; Sequence 2, Application US/09437895
; Patent No. 6416946
; GENERAL INFORMATION:
; APPLICANT: CHIEN, DAVID Y.
; ADDRESSEE: KUO, GEORGE
; TITLE OF INVENTION: METHODS OF TYPING HEPATITIS C VIRUS AND
; REAGENTS FOR USE THEREIN
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,553A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/060,400
; FILING DATE: 10-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: LEHNHARDT, SUSAN K.
; REGISTRATION NUMBER: 33,943
; REFERENCE/DOCKET NUMBER: 22300-20947.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 155 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-336-553A-2

Query Match          5.1%; Score 6; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      11 TSTWVL 16
Db      65 TSTWVL 70

RESULT 754
US-08-439-157-2
; Sequence 2, Application US/08439157
; Patent No. 6416944
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/437,895
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/336,553
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/060,400
FILING DATE: 10-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: LEHMHARDT, SUSAN K.
REGISTRATION NUMBER: 33,943
REFERENCE/DOCKET NUMBER: 22300-20947.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-437-895-2

Query Match 5.1%; Score 6; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TSTWVL 16
Db 65 TSTWVL 70

RESULT 756
US-09-252-991A-24143
Sequence 24143, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 24143
LENGTH: 157
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24143

Query Match 5.1%; Score 6; DB 2; Length 157;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 63 LAALAA 68

RESULT 757
US-09-252-991A-23802
Sequence 23802, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 23802
LENGTH: 161
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23802

Query Match 5.1%; Score 6; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 50 LAALAA 55

RESULT 758
US-09-673-395A-156
Sequence 156, Application US/09673395A
Patent No. 6620923
GENERAL INFORMATION:
APPLICANT: SPECHT, THOMAS
APPLICANT: HINZMANN, BERND
APPLICANT: SCHMITT, ARMIN
APPLICANT: PILARSKI, CHRISTIAN
APPLICANT: DAHL, EDGAR
APPLICANT: ROSENTHAL, ANDRE
TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE
CURRENT APPLICATION NUMBER: US/09/673,395A
CURRENT FILING DATE: 2000-10-17
NUMBER OF SEQ ID NOS: 637
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 156
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
US-09-673-395A-156

Query Match 5.1%; Score 6; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVL 55
Db 82 PDKEVL 87

RESULT 759
US-08-319-704-6
Sequence 6, Application US/08319704
Patent No. 5814617
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Charoenvit, Yupin
APPLICANT: Hedstrom, Richard C.
APPLICANT: Doolan, Denise L.
TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and
TITLE OF INVENTION: Erythrocytic Stage Immunogen and Gene
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Naval Medical R & D Command
STREET: Bldg 1, T-12, 8901 Wisconsin Avenue

; CITY: Bethesda
; STATE: Maryland
; COUNTRY: U.S.A
; ZIP: 20889-5606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/319,704
; FILING DATE: 07-OCT-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: A. David Spevack
; REGISTRATION NUMBER: 24,743
; REFERENCE/DOCKET NUMBER: 75,206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 295-6759
; TELEFAX: (301) 295-1022
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-319-704-6

Query Match 5.1%; Score 6; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
Db 91 VLLGGV 96

RESULT 760

US-09-605-703B-1008
; Sequence 1008, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1008
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-605-703B-1008

Query Match 5.1%; Score 6; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
Db 26 LLGGVL 31

RESULT 761

US-09-605-703B-1010
; Sequence 1010, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 60/142,764
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: 60/152,318
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1010
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-605-703B-1010

Query Match 5.1%; Score 6; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
Db 26 LLGGVL 31

RESULT 762

US-09-252-991A-29685
; Sequence 29685, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29685
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-29685

Query Match 5.1%; Score 6; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 FKGVKL 87
Db 28 FKGVKL 33

RESULT 763

US-09-583-110-4668
; Sequence 4668, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics

FILE REFERENCE: PATH00-07A
CURRENT APPLICATION NUMBER: US/09/593,110
CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 4668
LENGTH: 170
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-4668

Query Match 5.1%; Score 6; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 4.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 70 LAALAA 75

RESULT 764
US-09-071-035-442
Sequence 442, Application US/09071035
Patent No. 6448043
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 442:
SEQUENCE CHARACTERISTICS:
LENGTH: 172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-442

Query Match 5.1%; Score 6; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.8e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLL 21

Db 10 LGGVLL 15

RESULT 765
US-09-252-991A-26790
Sequence 26790, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26790
LENGTH: 172
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26790

Query Match 5.1%; Score 6; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.8e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQAT 94
DB 164 LLQAT 169

RESULT 766
US-10-206-576-442
Sequence 442, Application US/10206576
Patent No. 6913907
GENERAL INFORMATION:
APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369PID1
INFORMATION FOR SEQ ID NO: 442:

```
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 172 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 442:
US-10-206-576-442

Query Match          5.1%; Score 6; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLT 21
Db 10 LGGVLT 15

RESULT 767
US-08-956-171E-5218
; Sequence 5218, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 5218:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 173 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5218:
US-08-956-171E-5218

Query Match          5.1%; Score 6; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 22 AALAA 27
Db 165 AALAA 170

RESULT 768
US-08-781-986A-5218
; Sequence 5218, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunesch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 5218:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 173 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-781-986A-5218

Query Match          5.1%; Score 6; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
Db 165 AALAA 170

RESULT 769
US-09-902-540-14007
; Sequence 14007, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
```

; SEQ ID NO 14007
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14007

Query Match 5.1%; Score 6; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 10 LAALAA 15

RESULT 770

US-09-252-991A-30079
; Sequence 30079, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30079
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30079

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 AQVIAH 80
|||||
DB 32 AQVIAH 37

RESULT 771

US-09-134-000C-4177
; Sequence 4177, Application US/09134000C
; Patent No. 6617156

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4177
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-4177

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||||

DB 123 LLGGVL 128

RESULT 772

US-09-583-110-2827
; Sequence 2827, Application US/09583110
; Patent No. 6699703

GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 2827
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-2827

Query Match 5.1%; Score 6; DB 2; Length 186;

Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 77 VLAALA 82

RESULT 773

US-09-769-787-111
; Sequence 111, Application US/09769787
; Patent No. 6936252

GENERAL INFORMATION:

; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21123WO
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 111
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-111

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 77 VLAALA 82

RESULT 774

US-09-270-767-59012
; Sequence 59012, Application US/09270767
; Patent No. 6703491

; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 59012
; LENGTH: 187
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-59012

Query Match 5.1%; Score 6; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVT 11
Db 52 DLEVT 57

RESULT 775
US-09-489-039A-13345
; Sequence 13345, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13345
; LENGTH: 188
; TYPE: PRT
; ORGANISM: *Klebsiella pneumoniae*
US-09-489-039A-13345

Query Match 5.1%; Score 6; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 GLLORA 93
Db 181 GLLORA 186

RESULT 776
US-09-328-352-4549
; Sequence 4549, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4549
; LENGTH: 190
; TYPE: PRT
; ORGANISM: *Acinetobacter baumannii*
US-09-328-352-4549

Query Match 5.1%; Score 6; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
Db 98 GKVLGL 103

RESULT 777
US-09-134-000C-4429
; Sequence 4429, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4429
; LENGTH: 193
; TYPE: PRT
; ORGANISM: *Enterococcus faecalis*
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (11)..(11)
; OTHER INFORMATION: Amino acid 11 is Xaa wherein Xaa = any amino acid.
US-09-134-000C-4429

Query Match 5.1%; Score 6; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
Db 105 GKVLGL 110

RESULT 778
US-09-036-987A-23
; Sequence 23, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinoeyn Insecticide
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R
REGISTRATION NUMBER: 28,479
REFERENCE/DOCKET NUMBER: 50,608
TELEPHONE: (317)337-4816
TELEFAX: (317)337-4847
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 198 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-036-987A-23

Query Match 5.1%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LIGGVL 21
Db 15 LIGGVL 20

RESULT 779

US-09-370-700-23
Sequence 23, Application US/09370700
Patent No. 6274350
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 Div1
CURRENT APPLICATION NUMBER: US/09/370,700
CURRENT FILING DATE: 1999-08-09
EARLIER APPLICATION NUMBER: US 09/36987
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 198
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-370-700-23

Query Match 5.1%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LIGGVL 21
Db 15 LIGGVL 20

RESULT 780

US-09-603-207-23
Sequence 23, Application US/09603207B
Patent No. 6521406
GENERAL INFORMATION:
APPLICANT: Baltz, Richard H
APPLICANT: Broughton, Mary C
APPLICANT: Crawford, Kathryn P
APPLICANT: Madduri, Krishnamurthy
APPLICANT: Treadway, Patti J
APPLICANT: Turner, Jan R
APPLICANT: Waldron, Clive
TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
FILE REFERENCE: 50489 Div1
CURRENT APPLICATION NUMBER: US/09/603,207B

CURRENT FILING DATE: 2000-06-23
EARLIER APPLICATION NUMBER: 09/370,700
EARLIER FILING DATE: 1998-03-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 198
TYPE: PRT
ORGANISM: Saccharopolyspora spinosa
US-09-603-207-23

Query Match 5.1%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LIGGVL 21
Db 15 LIGGVL 20

RESULT 781

US-09-015-734-12
Sequence 12, Application US/09015734
Patent No. 6057127
GENERAL INFORMATION:
APPLICANT: Weber, Eric R.
APPLICANT: McCall, Catherine A.
TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/015,734
FILING DATE: 29-JAN-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: DI-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-09-015-734-12

Query Match 5.1%; Score 6; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LEVTTTS 12
Db 71 LEVTTTS 76

RESULT 782

US-09-515-311-12
Sequence 12, Application US/09515311

Patent No. 6582701
GENERAL INFORMATION:
APPLICANT: Weber, Eric R.
ATTORNEY/AGENT INFORMATION: McCall, Catherine A.
TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heeka Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/515,311
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/015,734
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: DI-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-515-311-12
Query Match 5.1%; Score 6; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 LEVTTS 12
Db 71 LEVTTS 76
RESULT 783
US-10-434-817-12
Sequence 12, Application US/10434817
Patent No. 6887672
GENERAL INFORMATION:
APPLICANT: Weber, Eric R.
McCall, Catherine A.
TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Heeka Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/515,311
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/015,734
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: DI-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-10-434-817-12
Query Match 5.1%; Score 6; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 LEVTTS 12
Db 71 LEVTTS 76
RESULT 784
US-09-540-236-2414
Sequence 2414, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 2414
LENGTH: 202
TYPE: PRT
ORGANISM: M.catarrhalis
US-09-540-236-2414
Query Match 5.1%; Score 6; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 84 GKVLGL 89
Db 106 GKVLGL 111
RESULT 785
US-09-107-433-3144
Sequence 3144, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A. Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts

; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085111
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3144:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...202
; SEQUENCE DESCRIPTION: SEQ ID NO: 3144:
US-09-107-433-3144

Query Match 5.1%; Score 6; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
DB 93 VLAALA 98

RESULT 786
US-09-252-991A-24035
; Sequence 24035, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24035
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24035

Query Match 5.1%; Score 6; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

DB 147 VLAALA 152

RESULT 787
US-09-257-583-15
; Sequence 15, Application US/09257583A
; Patent No. 6429362
; GENERAL INFORMATION:
; APPLICANT: Crane, Virginia
; TITLE OF INVENTION: Family Of Maize PR-1 Genes And Promoters
; FILE REFERENCE: 5718-32, 035718/175219
; CURRENT APPLICATION NUMBER: US/09/257,583A
; CURRENT FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Zea mays
US-09-257-583-15

Query Match 5.1%; Score 6; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 10 LAALAA 15

RESULT 788
US-09-252-991A-19931
; Sequence 19931, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19931
; LENGTH: 218
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19931

Query Match 5.1%; Score 6; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 169 LLGGVL 174

RESULT 789
US-09-482-273-119
; Sequence 119, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849

```
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 119
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (51)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-482-273-119

Query Match          5.1%; Score 6; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 5 LAALAA 10

RESULT 790
US-09-583-110-5016
; Sequence 5016, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5016
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5016

Query Match          5.1%; Score 6; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 18 VLAALA 23

RESULT 791
US-09-769-787-86
; Sequence 86, Application US/09769787
; Patent No. 6936252
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics Limited
; APPLICANT: Gilbert, Christophe FG
; APPLICANT: Hanbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21129W0
; CURRENT APPLICATION NUMBER: US/09/769,787
; CURRENT FILING DATE: 2001-01-26
```

```
; PRIOR APPLICATION NUMBER: GB 9816337.1
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: US 60/125164
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 388
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 86
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-787-86

Query Match          5.1%; Score 6; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 18 VLAALA 23

RESULT 792
US-08-886-765-2
; Sequence 2, Application US/08886765
; Patent No. 5817500
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter Kamp
; APPLICANT: Wagner, Peter
; APPLICANT: Mullertz, Anette
; APPLICANT: Knap, Inge Helmer
; TITLE OF INVENTION: Animal Feed Additives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: No. 5817500 No. 5817500disk of No. 5817500th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/886,765
; FILING DATE: 1-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4324.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-886-765-2

Query Match          5.1%; Score 6; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 9 LAALAA 14

RESULT 793
```


US-09-115-660-2
; Sequence 2, Application US/09115660
; Patent No. 6245546
; GENERAL INFORMATION:
; APPLICANT: Hansen, Peter Kamp
; APPLICANT: Wagner, Peter
; APPLICANT: Mullertz, Anette
; APPLICANT: Knap, Inge Helmer
; TITLE OF INVENTION: Animal Feed Additives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6245546 No. 6245546disk of No. 6245546th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/115,660
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,765
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4324.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 225 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-115-660-2
Query Match 5.1%; Score 6; DB 2; Length 225;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 LAALAA 26
DB 9 LAALAA 14
RESULT 794
US-09-252-991A-17728
; Sequence 17728, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17728
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17728

Query Match 5.1%; Score 6; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 LAALAA 26
DB 155 LAALAA 160
RESULT 795
US-09-071-035-476
; Sequence 476, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 476:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 227 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-476

Query Match 5.1%; Score 6; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 APYIEQ 74
DB 144 APYIEQ 149

RESULT 796

US-10-206-576-476
; Sequence 476, Application US/10206576
; Patent No. 6913907
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-R
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14
ATTORNEY/AGENT INFORMATION:
NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369P1D1
INFORMATION FOR SEQ ID NO: 476:
SEQUENCE CHARACTERISTICS:
LENGTH: 227 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 476:
US-10-206-576-476

Query Match 5.1%; Score 6; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 APYIEQ 74
Db 144 APYIEQ 149

RESULT 797
US-09-605-703B-818
Sequence 818, Application US/09605703B
Patent No. 6962989
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
FILE REFERENCE: BGI-129CP
CURRENT APPLICATION NUMBER: US/09/605,703B
CURRENT FILING DATE: 2000-06-27
PRIOR APPLICATION NUMBER: 60/142,764
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: 60/152,318
PRIOR FILING DATE: 1999-09-03
NUMBER OF SEQ ID NOS: 2934
SEQ ID NO 818
LENGTH: 227
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-605-703B-818
Query Match 5.1%; Score 6; DB 2; Length 227;

Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 101 EPIVTT 106
Db 80 EPIVTT 85

RESULT 798
US-09-252-991A-30629
Sequence 30629, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 30629
LENGTH: 229
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30629

Query Match 5.1%; Score 6; DB 2; Length 229;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 LAALAA 26
Db 60 LAALAA 65

RESULT 799
US-09-107-532A-7215
Sequence 7215, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:

TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 7215:
SEQUENCE CHARACTERISTICS:
LENGTH: 230 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...230
SEQUENCE DESCRIPTION: SEQ ID NO: 7215:
US-09-107-532A-7215

Query Match 5.1%; Score 6; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 IEQAQV 77
|||||
Db 110 IEQAQV 115

RESULT 800
US-09-328-352-7928
; Sequence 7928, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7928
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7928

Query Match 5.1%; Score 6; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 112 GVLAAL 117

RESULT 801
US-09-949-016-7831
; Sequence 7831, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7831

LENGTH: 231
TYPE: PRT
ORGANISM: Human
US-09-949-016-7831

Query Match 5.1%; Score 6; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQR 92
|||||
Db 3 LGLLQR 8

RESULT 802
US-09-270-767-42502
; Sequence 42502, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42502
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-42502

Query Match 5.1%; Score 6; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 47 GVLAAL 52

RESULT 803
US-09-107-433-5113
; Sequence 5113, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5113:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...233
SEQUENCE DESCRIPTION: SEQ ID NO: 5113:
US-09-107-433-5113

Query Match 5.1%; Score 6; DB 2; Length 233;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 30 VLAALA 35

RESULT 804
US-09-015-734-7
; Sequence 7, Application US/09015734
; Patent No. 6057127
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-015-734-7

Query Match 5.1%; Score 6; DB 2; Length 236;

Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 7 LEVTTTS 12
Db 52 LEVTTTS 57

RESULT 805
US-09-515-311-7
; Sequence 7, Application US/09515311
; Patent No. 6582701
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/515,311
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/015,734
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
US-09-515-311-7

Query Match 5.1%; Score 6; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LEVTTTS 12
Db 52 LEVTTTS 57

RESULT 806
US-10-434-817-7
; Sequence 7, Application US/10434817
; Patent No. 6887672
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation

STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/434,817
FILING DATE: 08-May-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/015,734
FILING DATE: 29-JAN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: DI-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-434-817-7

Query Match 5.1%; Score 6; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LEVTTTS 12
DB 52 LEVTTTS 57

RESULT 807
US-09-248-796A-18049
; Sequence 18049, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18049
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18049

Query Match 5.1%; Score 6; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 IVPDKK 53
DB 198 IVPDKK 203

RESULT 808
US-09-252-991A-23387
; Sequence 23387, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23387
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23387

Query Match 5.1%; Score 6; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQR 92
DB 32 LGLLQR 37

RESULT 809
US-09-902-540-12387
; Sequence 12387, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12387
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12387

Query Match 5.1%; Score 6; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 84 LAALAA 89

RESULT 810
US-09-252-991A-28112
; Sequence 28112, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

;; PRIOR APPLICATION NUMBER: US 60/074,788
;; PRIOR FILING DATE: 1998-02-18
;; PRIOR APPLICATION NUMBER: US 60/094,190
;; PRIOR FILING DATE: 1998-07-27
;; NUMBER OF SEQ ID NOS: 33142
;; SEQ ID NO 28112
;; LENGTH: 240
;; TYPE: PRT
;; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28112

Query Match 5.1%; Score 6; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 CSQAAP 70
| | | | |
Db 130 CSQAAP 135

RESULT 811

US-09-252-991A-19639
; Sequence 19639, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19639
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19639

Query Match 5.1%; Score 6; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 91 QRATQQ 96
| | | | |
Db 20 QRATQQ 25

RESULT 812

US-09-252-991A-27611
; Sequence 27611, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27611
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27611

Query Match 5.1%; Score 6; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 VIEPIV 104
| | | | |
Db 113 VIEPIV 118

RESULT 813

US-09-252-991A-24362
; Sequence 24362, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24362
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24362

Query Match 5.1%; Score 6; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 70 PYIEQA 75
| | | | |
Db 83 PYIEQA 88

RESULT 814

US-09-015-734-2
; Sequence 2, Application US/09015734
; Patent No. 6057127
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; TITLE OF INVENTION: NOVEL EQUINE FC EPSILON RECEPTOR ALPHA
; TITLE OF INVENTION: CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/015,734
; FILING DATE: 29-JAN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: DI-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-09-015-734-2

Query Match 5.1%; Score 6; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LEVITS 12
|||||
DB 71 LEVITS 76

RESULT 815

US-09-515-311-2
Sequence 2, Application US/09515311

Patent No. 6582701

GENERAL INFORMATION:

APPLICANT: Weber, Eric R.

ADDRESSEE: Heeska Corporation

TITLE OF INVENTION: NOVEL EQUINE PC EPSILON RECEPTOR ALPHA

CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heeska Corporation

STREET: 1825 Sharp Point Drive

CITY: Fort Collins

STATE: Colorado

COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/515,311

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/015,734

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: DI-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 255 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: Protein

US-09-515-311-2

Query Match 5.1%; Score 6; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LEVITS 12
|||||
DB 71 LEVITS 76

RESULT 816

US-09-902-540-13380

Sequence 13380, Application US/09902540

Patent No. 6833447

GENERAL INFORMATION:

APPLICANT: Goldman, Barry S.

ADDRESSEE: Hinkle, Gregory J.

ADDRESSEE: Slater, Steven C.

ADDRESSEE: Wiegand, Roger C.

TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

FILE REFERENCE: 38-10(15849)B

CURRENT APPLICATION NUMBER: US/09/902,540

CURRENT FILING DATE: 2001-07-10

PRIOR APPLICATION NUMBER: 60/217,883

PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825

SEQ ID NO 13380

LENGTH: 255

TYPE: PRT

ORGANISM: Myxococcus xanthus

US-09-902-540-13380

Query Match 5.1%; Score 6; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAY 27
|||||
DB 193 AALAY 198

RESULT 817

US-10-434-817-2

Sequence 2, Application US/10434817

Patent No. 6887672

GENERAL INFORMATION:

APPLICANT: Weber, Eric R.

ADDRESSEE: McCall, Catherine A.

TITLE OF INVENTION: NOVEL EQUINE PC EPSILON RECEPTOR ALPHA

CHAIN NUCLEIC ACID MOLECULES, PROTEINS AND USES THEREOF

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Heeska Corporation

STREET: 1825 Sharp Point Drive

CITY: Fort Collins

STATE: Colorado

COUNTRY: USA

ZIP: 80525

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WordPerfect for Windows, Version 7.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/434,817

FILING DATE: 08-May-2003

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/015,734

FILING DATE: 29-JAN-1998

ATTORNEY/AGENT INFORMATION:

NAME: Verser, Carol Talkington

REGISTRATION NUMBER: 37,459

REFERENCE/DOCKET NUMBER: DI-4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 970/493-7272

TELEFAX: 970/484-9505

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 255 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: Protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2;

US-10-434-817-2

```
Query Match          5.1%; Score 6; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LEVITS 12
   |||||
Db 71 LEVITS 76

RESULT 818
US-09-252-991A-20260
; Sequence 20260, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20260
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20260

Query Match          5.1%; Score 6; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
   |||||
Db 174 LGGVLA 179

RESULT 819
US-09-252-991A-20609
; Sequence 20609, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20609
; LENGTH: 261
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20609

Query Match          5.1%; Score 6; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
   |||||
Db 209 LGGVLA 214

RESULT 820
US-09-252-991A-32324
```

```
; Sequence 32324, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: MARC J. RUBENFIELD ET AL.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32324
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32324

Query Match          5.1%; Score 6; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
   |||||
Db 260 VLAALA 265

RESULT 821
US-08-311-731A-114
; Sequence 114, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 114:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-114
```


Query Match 5.1%; Score 6; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRAT 94
|||||
DB 203 LLQRAT 208

RESULT 822
US-09-813-453B-15
; Sequence 15, Application US/09813453B
; Patent No. 6830898
; GENERAL INFORMATION:
; APPLICANT: Yocum, R. Rogers
; TITLE OF INVENTION: MICROORGANISMS AND ASSAYS FOR THE IDENTIFICATION OF
; FILE REFERENCE: CGZ-001
; CURRENT APPLICATION NUMBER: US/09/813,453B
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/227,860
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/667,569
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Bordetella pertussis
US-09-813-453B-15

Query Match 5.1%; Score 6; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 256 LAALAA 261

RESULT 823
US-09-270-767-32500
; Sequence 32500, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32500
; LENGTH: 269
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-32500

Query Match 5.1%; Score 6; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
|||||
DB 188 KVLGLL 193

RESULT 824
US-09-252-991A-18538
; Sequence 18538, Application US/09252991A

Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18538
; LENGTH: 271
; TYPE: PRT
; ORGANISM: *Pseudomonas aeruginosa*
US-09-252-991A-18538

Query Match 5.1%; Score 6; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
DB 194 VLLGGV 199

RESULT 825
US-09-489-039A-8452
; Sequence 8452, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8452
; LENGTH: 275
; TYPE: PRT
; ORGANISM: *Klebsiella pneumoniae*
US-09-489-039A-8452

Query Match 5.1%; Score 6; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQ 58
|||||
DB 84 EVLYQQ 89

RESULT 826
US-09-583-110-5207
; Sequence 5207, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to *Streptococcus*
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553

; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5207
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-593-110-5207

Query Match 5.1%; Score 6; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 25 GGVLA 30

RESULT 827
US-09-489-039A-12241
; Sequence 12241, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12241
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12241

Query Match 5.1%; Score 6; DB 2; Length 282;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 182 LAALAA 187

RESULT 828
US-09-489-039A-9220
; Sequence 9220, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9220
; LENGTH: 283
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9220

Query Match 5.1%; Score 6; DB 2; Length 283;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 209 LGGVLA 214

RESULT 829
US-09-252-991A-20552
; Sequence 20552, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20552
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (219)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20552

Query Match 5.1%; Score 6; DB 2; Length 285;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 264 GGVLA 269

RESULT 830
US-09-328-352-7096
; Sequence 7096, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328.352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7096
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7096

Query Match 5.1%; Score 6; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 QAVIEP 102
Db 234 QAVIEP 239

RESULT 831
US-09-540-236-2477
; Sequence 2477, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001

; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2477
; LENGTH: 287
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2477

Query Match 5.1%; Score 6; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EQAQVI 78
Db 120 EQAQVI 125

RESULT 832
US-09-902-540-13218
; Sequence 13218, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(115849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13218
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13218

Query Match 5.1%; Score 6; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 150 LGGVLA 155

RESULT 833
US-09-252-991A-27979
; Sequence 27979, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27979
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27979

Query Match 5.1%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 89 LLQRAT 94
Db 240 LLQRAT 245

RESULT 834
US-09-949-016-11675
; Sequence 11675, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11675
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11675

Query Match 5.1%; Score 6; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRA 93
Db 18 GLLQRA 23

RESULT 835
US-09-071-035-480
; Sequence 480, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 480:
SEQUENCE CHARACTERISTICS:
LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-480

Query Match 5.1%; Score 6; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 APYIEQ 74
|||||
DB 57 APYIEQ 62

RESULT 836

US-10-206-576-480
Sequence 480, Application US/10206576
Patent No. 6913907
GENERAL INFORMATION:

APPLICANT: Choi et al.
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 497
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: CD-R
COMPUTER: Dell Latitude
OPERATING SYSTEM: Windows 98
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/206,576
FILING DATE: 29-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/071,035
FILING DATE: 1998-05-04
APPLICATION NUMBER: US 60/046,655
FILING DATE: 1997-05-16
APPLICATION NUMBER: US 60/044,031
FILING DATE: 1997-05-06
APPLICATION NUMBER: US 60/066,009
FILING DATE: 1997-11-14

ATTORNEY/AGENT INFORMATION:

NAME: Hyman, Mark J.
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: PB369PID1
INFORMATION FOR SEQ ID NO: 480:

SEQUENCE CHARACTERISTICS:

LENGTH: 289 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 480:

Query Match 5.1%; Score 6; DB 2; Length 289;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 APYIEQ 74
|||||

Db 57 APYIEQ 62

RESULT 837

US-09-328-352-7390
Sequence 7390, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 7390
LENGTH: 290
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-7390

Query Match 5.1%; Score 6; DB 2; Length 290;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 QAQVIA 79
|||||
DB 69 QAQVIA 74

RESULT 838

US-09-252-991A-22826
Sequence 22826, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 22826
LENGTH: 291
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22826

Query Match 5.1%; Score 6; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAY 27
|||||
DB 269 AALAAAY 274

RESULT 839

US-09-724-623-103
Sequence 103, Application US/09724623
Patent No. 6476209
GENERAL INFORMATION:

APPLICANT: Glenn, Matthew
APPLICANT: Lubbers, Mark W
APPLICANT: Dekker, James
TITLE OF INVENTION: Polynucleotides, materials incorporating
TITLE OF INVENTION: them, and methods for using them.
FILE REFERENCE: 1048UI
CURRENT APPLICATION NUMBER: US/09/724,623

; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Lactobacillus rhamnosus
US-09-724-623-103

Query Match 5.1%; Score 6; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 174 LGGVLA 179

RESULT 840
US-08-811-481-4
; Sequence 4, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal

Query Match 5.1%; Score 6; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWVIV 37
|||||

DB 199 GCWVIV 204

RESULT 841
US-09-876-527-4
; Sequence 4, Application US/09876527
; Patent No. 6627735
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,527
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/811,481
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-876-527-4

Query Match 5.1%; Score 6; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWVIV 37
|||||
DB 199 GCWVIV 204

RESULT 842
US-09-252-991A-22835
; Sequence 22835, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18 DB 2; Length 299;
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22835
; LENGTH: 299
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22835

Query Match 5.1%; Score 6; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
| | | | |
Db 210 GKVLGL 215

RESULT 843
US-09-252-991A-26364
; Sequence 26364, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26364
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26364

Query Match 5.1%; Score 6; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
| | | | |
Db 237 LGGVLA 242

RESULT 844
US-09-252-991A-24026
; Sequence 24026, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24026
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24026

Query Match 5.1%; Score 6; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 237 VLAALA 242

RESULT 845
US-09-328-352-5720
; Sequence 5720, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5720
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5720

Query Match 5.1%; Score 6; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
| | | | |
Db 48 LLGGVL 53

RESULT 846
US-09-328-352-6103
; Sequence 6103, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6103
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6103

Query Match 5.1%; Score 6; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
| | | | |
Db 281 VLLGGV 286

RESULT 847
US-09-248-796A-20160
; Sequence 20160, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20160
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20160

Query Match 5.1%; Score 6; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 238 LAALAA 243

RESULT 848

US-09-160-975A-2
; Sequence 2, Application US/09160975A
; Patent No. 6406883
; GENERAL INFORMATION:

; APPLICANT: Barbara Spellerberg, Rudolf L t t i c k e n, Andreas
; APPLICANT: Podbielski, and Eva Rozdzinski
; TITLE OF INVENTION: LMB, POLYNUCLEOTIDE, POLYPEPTIDES
; TITLE OF INVENTION: AND METHOD OF USE THEREFOR
; NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: PENTIUM
; OPERATING SYSTEM: WINDOWS 95

SOFTWARE: ASCII

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/160,975A
; FILING DATE: Concurrently

CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: LILLIE, Raymond J.

; REGISTRATION NUMBER: 31,778

; REFERENCE/DOCKET NUMBER:

; TELEPHONE: 973-994-1700

; TELEFAX: 973-994-1744

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 306 AMINO ACIDS

; TYPE: AMINO ACID

; TOPOLOGY: LINEAR

; MOLECULE TYPE: PROTEIN

US-09-160-975A-2

Query Match 5.1%; Score 6; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQ 58
|||||
DB 299 EVLYQQ 304

RESULT 849

US-09-252-991A-26798
; Sequence 26798, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26798

; LENGTH: 306

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26798

Query Match

Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 ATQOQA 98
|||||
DB 79 ATQOQA 84

RESULT 850

US-09-248-796A-20451
; Sequence 20451, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 20451

; LENGTH: 316

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-20451

Query Match

Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MSADLE 8
|||||
DB 17 MSADLE 22

RESULT 851

US-09-252-991A-23561
; Sequence 23561, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 33561
; LENGTH: 319
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23561

Query Match 5.1%; Score 6; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 180 LAALAA 185

RESULT 852
US-09-583-110-3120
; Sequence 3120, Application US/09583110
; Patent No. 6899703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 3120
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-3120

Query Match 5.1%; Score 6; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
| | | | |
Db 160 GGVLA 165

RESULT 853
US-09-902-540-13143
; Sequence 13143, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 2000-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13143
; LENGTH: 321

; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13143

Query Match 5.1%; Score 6; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 137 VLAALA 142

RESULT 854
US-09-107-433-3055
; Sequence 3055, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154

; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 3055:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 322 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8) LOCATION 1...322
; SEQUENCE DESCRIPTION: SEQ ID NO: 3055:
US-09-107-433-3055

Query Match 5.1%; Score 6; DB 2; Length 322;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
| | | | |
Db 161 GGVLA 166

RESULT 855

US-09-902-540-12351
; Sequence 12351, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12351
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12351

Query Match 5.1%; Score 6; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LIGGYL 21
Db 217 LIGGYL 222

RESULT 856

US-10-272-490-42
; Sequence 42, Application US/10272490
; Patent No. 6943001
; GENERAL INFORMATION:
; APPLICANT: Zhao, Lishan
; APPLICANT: Mathur, Eric J.
; APPLICANT: Weiner, David
; APPLICANT: Richardson, Toby
; APPLICANT: Milan, Aileen
; APPLICANT: Burk, Mark J.
; APPLICANT: Han, Bin
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: EXPOXIDE HYDROLASES, NUCLEIC ACIDS ENCODING THEM AND METHODS
; FILE REFERENCE: 09010-831001
; CURRENT APPLICATION NUMBER: US/10/272,490
; CURRENT FILING DATE: 2002-10-10
; PRIOR APPLICATION NUMBER: US 10/214,473
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US 60/309,478
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/393,378
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental source
US-10-272-490-42

Query Match 5.1%; Score 6; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 48 LAALAA 53

RESULT 857

US-09-252-991A-27058
; Sequence 27058, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27058
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27058

Query Match 5.1%; Score 6; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 246 VLAALA 251

RESULT 858

US-09-248-796A-24407
; Sequence 24407, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24407
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24407

Query Match 5.1%; Score 6; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 EPIVTT 106
Db 57 EPIVTT 62

RESULT 859

US-09-724-797-28
; Sequence 28, Application US/09724797
; Patent No. 673398
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24407
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24407

Query Match 5.1%; Score 6; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 48 LAALAA 53

```

; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 2000-11-28
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Bacteria
US-09-724-797-28

Query Match          5.1%; Score 6; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
Db 219 LLGGVL 224

RESULT 860
US-09-252-991A-26165
; Sequence 26165, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26165
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26165

Query Match          5.1%; Score 6; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGLLQ 91
Db 57 VLGLLQ 62

RESULT 861
US-09-902-540-9720
; Sequence 9720, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 9720
; LENGTH: 332
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

```

```

US-09-902-540-9720

Query Match          5.1%; Score 6; DB 2; Length 332;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 286 LAALAA 291

RESULT 862
US-09-758-759-204
; Sequence 204, Application US/09758759
; Patent No. 6861513
; GENERAL INFORMATION:
; APPLICANT: Hosted, Thomas J.
; APPLICANT: Wang, Tim X.
; APPLICANT: Horan, Ann C.
; TITLE OF INVENTION: Evernimycin Biosynthetic Genes
; FILE REFERENCE: ID0983K US
; CURRENT APPLICATION NUMBER: US/09/758,759
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/175,751
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 204
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Micromonospora carbonacea
; FEATURE:
; OTHER INFORMATION: ORF11
US-09-758-759-204

Query Match          5.1%; Score 6; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 1 VLAALA 6

RESULT 863
US-09-570-856B-15
; Sequence 15, Application US/09570856B
; Patent No. 6682923
; GENERAL INFORMATION:
; APPLICANT: Bentzien, Joerg M
; APPLICANT: Dahiyat, Basil I
; TITLE OF INVENTION: NOVEL THERMOSTABLE ALKALIPHILIC XYLANASE
; FILE REFERENCE: A-67478-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/570,856B
; CURRENT FILING DATE: 2002-04-15
; PRIOR APPLICATION NUMBER: US 60/133,714
; PRIOR FILING DATE: 1999-05-12
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Streptomyces lividans
US-09-570-856B-15

Query Match          5.1%; Score 6; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 27 LAALAA 32

```

RESULT 864
US-09-489-039A-10149
; Sequence 10149, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10149
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10149

Query Match 5.1%; Score 6; DB 2; Length 337;
Best Local Similarity 100.0%; Pred. No. 8.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
| | | | |
DB 44 LAALAA 49

RESULT 865
US-09-252-991A-24373
; Sequence 24373, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24373
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; NAME/KEY: UNSURE
; LOCATION: (98)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-24373

Query Match 5.1%; Score 6; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 8.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
| | | | |
DB 81 VLAALA 86

RESULT 866
US-09-252-991A-27223
; Sequence 27223, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27223
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27223

Query Match 5.1%; Score 6; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
| | | | |
DB 264 LAALAA 269

RESULT 867
US-09-489-039A-8662
; Sequence 8662, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8662
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8662

Query Match 5.1%; Score 6; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
| | | | |
DB 113 LAALAA 118

RESULT 868
US-09-543-681A-4974
; Sequence 4974, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4974
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4974

Query Match 5.1%; Score 6; DB 2; Length 346;

Best Local Similarity 100.0%; Pred. No. 9.1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
Db 210 AALAA 215
|||||

RESULT 869
US-09-248-796A-16397
; Sequence 16397, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16397
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16397

Query Match 5.1%; Score 6; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 IEQAQV 77
Db 251 IEQAQV 256
|||||

RESULT 870
US-09-489-039A-9489
; Sequence 9489, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9489
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9489

Query Match 5.1%; Score 6; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVL 21
Db 241 LGGVL 246
|||||

RESULT 871
US-09-252-991A-20858
; Sequence 20858, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20858
; LENGTH: 349
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20858

Query Match 5.1%; Score 6; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 66 VLAALA 71
|||||

RESULT 872
US-09-489-039A-10384
; Sequence 10384, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10384
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10384

Query Match 5.1%; Score 6; DB 2; Length 350;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 225 GVLAAL 230
|||||

RESULT 873
US-09-489-039A-13104
; Sequence 13104, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13104
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13104

US-09-489-039A-13104

Query Match 5.1%; Score 6; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21

DB 132 LLGGVL 137

RESULT 874

US-09-602-787A-466

; Sequence 466, Application US/09602787A

; Patent No. 6696561

; GENERAL INFORMATION:

; APPLICANT: Pompejus, Mark

; APPLICANT: Krüger, Burkhard

; APPLICANT: Schöder, Hartwig

; APPLICANT: Zelder, Oskar

; APPLICANT: Habernauer, Gregor

; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS

; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE

; TITLE OF INVENTION: TRANSPORT

; FILE REFERENCE: BGI-125CP

; CURRENT APPLICATION NUMBER: US/09/602,787A

; CURRENT FILING DATE: 2000-06-23

; PRIOR APPLICATION NUMBER: USSN 60/141031

; PRIOR FILING DATE: 1999-06-25

; PRIOR APPLICATION NUMBER: DE 19931454.3

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931478.0

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19931563.9

; PRIOR FILING DATE: 1999-07-08

; PRIOR APPLICATION NUMBER: DE 19932122.1

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932124.8

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932125.6

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932128.0

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932180.9

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932182.5

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932190.6

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932191.4

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932209.0

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932212.0

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932227.9

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932228.7

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932229.5

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932230.9

; PRIOR FILING DATE: 1999-07-09

; PRIOR APPLICATION NUMBER: DE 19932927.3

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19933005.0

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19933006.9

; PRIOR FILING DATE: 1999-07-14

; PRIOR APPLICATION NUMBER: DE 19940764.9

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19940765.7

; PRIOR FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: DE 19940766.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940830.0
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940831.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940832.7
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19940833.5
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941395.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942077.7
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942078.5
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942079.3
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 678
; SEQ ID NO 466
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-602-787A-466

Query Match 5.1%; Score 6; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26

DB 34 LAALAA 39

RESULT 875

US-09-542-767A-1

; Sequence 1, Application US/09542767A

; Patent No. 6296671

; GENERAL INFORMATION:

; APPLICANT: Schuelein, Martin

; APPLICANT: Kristensen, Henrik

; TITLE OF INVENTION: An Enzymatic Treatment Method

; FILE REFERENCE: 5871.204-US

; CURRENT APPLICATION NUMBER: US/09/542,767A

; CURRENT FILING DATE: 2000-04-04

; PRIOR APPLICATION NUMBER: PA 1999 00390

; PRIOR FILING DATE: 1999-03-22

; PRIOR APPLICATION NUMBER: US 60/125,884

; PRIOR FILING DATE: 1999-03-24

; PRIOR APPLICATION NUMBER: PCT/DK00/00136

; PRIOR FILING DATE: 2000-03-22

; NUMBER OF SEQ ID NOS: 1

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 360

; TYPE: PRT

; ORGANISM: Trichisporon penicillatum

US-09-542-767A-1

Query Match

5.1%; Score 6; DB 2; Length 360;

Best Local Similarity 100.0%; Pred. No. 9.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26

DB 12 LAALAA 17

RESULT 876
US-09-252-991A-29984
; Sequence 29984, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29984
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29984
Query Match 5.1%; Score 6; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 74 QAQVIA 79
Db 346 QAQVIA 351
|||||
RESULT 877
US-09-198-452A-604
; Sequence 604, Application US/09198452A
; Patent No. 655294
; GENERAL INFORMATION:
; APPLICANT: Grifais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 604
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; NAME/KEY: SITE
; LOCATION: 1...367
; OTHER INFORMATION: Xaa=unknown or other
US-09-198-452A-604
Query Match 5.1%; Score 6; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 LGGVLA 22
Db 41 LGGVLA 46
|||||
RESULT 878
US-09-794-534-2
; Sequence 2, Application US/09794534
; Patent No. 6664083
; GENERAL INFORMATION:
; APPLICANT: Matsuyama, Akinobu
; APPLICANT: Tokuyama, Shinji
; TITLE OF INVENTION: Methods for racemizing N-acylamino acids and
; TITLE OF INVENTION: producing optically active amino acids

; FILE REFERENCE: 06501-074001
; CURRENT APPLICATION NUMBER: US/09/794,534
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: JP 2000-060358
; PRIOR FILING DATE: 2000-03-01
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Sebekia benihana
US-09-794-534-2
Query Match 5.1%; Score 6; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 8 EVTTST 13
Db 356 EVTTST 361
|||||
RESULT 879
US-09-252-991A-20724
; Sequence 20724, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20724
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20724
Query Match 5.1%; Score 6; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 LAALAA 26
Db 39 LAALAA 44
|||||
RESULT 880
US-09-252-991A-20081
; Sequence 20081, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20081
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

FEATURE:
NAME/KEY: UNSURE
LOCATION: (201), (226), (315)
OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20081

Query Match 5.1%; Score 6; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 ATQOQA 98
|||||

Db 3 ATQOQA 8
|||||

RESULT 881

US-09-826-509-457
Sequence 457, Application US/09826509
Patent No. 6806054

GENERAL INFORMATION:
APPLICANT: Lehmann-Bruinsma, Karin
APPLICANT: Liaw, Chen W.

APPLICANT: Lin, I-Lin

TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G

TITLE OF INVENTION: Protein-Coupled Receptors

FILE REFERENCE: AREN-207

CURRENT APPLICATION NUMBER: US/09/826,509

PRIOR FILING DATE: 2001-04-05

PRIOR APPLICATION NUMBER: 60/195,747

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: 09/170,496

PRIOR FILING DATE: 1998-10-13

NUMBER OF SEQ ID NOS: 589

SOFTWARE: PatentIn Version 2.1

SEQ ID NO 457

LENGTH: 371

TYPE: PRT

ORGANISM: Homo sapiens

US-09-826-509-457

Query Match 5.1%; Score 6; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||

Db 58 VLAALA 63
|||||

RESULT 882

US-09-252-991A-31788

Sequence 31788, Application US/09252991A

Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 31788

LENGTH: 372

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-31788

Query Match 5.1%; Score 6; DB 2; Length 372;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||||

Db 251 LLGGVL 256
|||||

RESULT 883

US-09-740-288A-12

Sequence 12, Application US/09740288A

Patent No. 6849783

GENERAL INFORMATION:

APPLICANT: Allen, Stephen

APPLICANT: Kinney, Anthony

APPLICANT: Miao, Guo-Hua

APPLICANT: Orozco, Emil

TITLE OF INVENTION: PLANT BIOTIN SYNTHASE

FILE REFERENCE: BB1429 US NA

CURRENT APPLICATION NUMBER: US/09/740,288A

CURRENT FILING DATE: 2000-12-19

PRIOR APPLICATION NUMBER: US 60/172929

PRIOR FILING DATE: 1999-12-21

NUMBER OF SEQ ID NOS: 36

SOFTWARE: Microsoft Office 97

SEQ ID NO 12

LENGTH: 374

TYPE: PRT

ORGANISM: Glycine max

US-09-740-288A-12

Query Match 5.1%; Score 6; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
|||||

Db 345 KVLGLL 350
|||||

RESULT 884

US-08-883-534-5

Sequence 5, Application US/08883534

Patent No. 5846777

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Lal, Preeti

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: TWO NEW WD-40 PROTEINS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/883,534

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0332 US

TELECOMMUNICATION INFORMATION:

```
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 376 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 1314316
US-08-883-534-5

Query Match 5.1%; Score 6; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LEVTTTS 12
Db 46 LEVTTTS 51

RESULT 885
US-09-204-764-5
; Sequence 5, Application US/09204764
; Patent No. 6025464
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/09/204,764
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/983,534
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0332 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1314316
US-09-204-764-5

Query Match 5.1%; Score 6; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LEVTTTS 12
Db 46 LEVTTTS 51

RESULT 886
US-08-811-481-7
; Sequence 7, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/811,481
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-811-481-7

Query Match 5.1%; Score 6; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIV 37
Db 199 GCVVIV 204

RESULT 887
US-09-876-527-7
; Sequence 7, Application US/09876527
; Patent No. 6627735
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
```

```
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 5:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 376 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 1314316
US-08-883-534-5

Query Match 5.1%; Score 6; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LEVTTTS 12
Db 46 LEVTTTS 51

RESULT 886
US-08-811-481-7
; Sequence 7, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/811,481
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-811-481-7

Query Match 5.1%; Score 6; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIV 37
Db 199 GCVVIV 204

RESULT 887
US-09-876-527-7
; Sequence 7, Application US/09876527
; Patent No. 6627735
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
```


;; Hagopian, William A.
;; Lacasse, James M.
;; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
;; NUMBER OF SEQUENCES: 34
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: ZymoGenetics, Inc.
;; STREET: 1201 Eastlake Avenue East
;; CITY: Seattle
;; STATE: WA
;; COUNTRY: USA
;; ZIP: 98102
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; FILING DATE: 07-Jun-2001
;; CLASSIFICATION: <Unknown>
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/811,481
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Lingenfelter, Susan
;; REGISTRATION NUMBER: P-41,156
;; REFERENCE/DOCKET NUMBER: 95-36
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 206-442-6675
;; TELEFAX: 206-442-6678
;; TELEX: <Unknown>
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 376 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FRAGMENT TYPE: internal
;; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-876-527-7

Query Match 5.1%; Score 6; DB 2; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWIV 37
Db 199 GCWIV 204

RESULT 888
US-09-107-532A-5519
; Sequence 5519, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/107,532A
;; FILING DATE: 30-Jun-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/085,598
;; FILING DATE: 14 May 1998
;; APPLICATION NUMBER: 60/051571
;; FILING DATE: July 2, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ariniello, Pamela Deneka
;; REGISTRATION NUMBER: 40,489
;; REFERENCE/DOCKET NUMBER: GTC-012
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781)893-5007
;; TELEFAX: (781)893-8277
;; INFORMATION FOR SEQ ID NO: 5519:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 379 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Enterococcus faecium
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (B) LOCATION 1...379
;; SEQUENCE DESCRIPTION: SEQ ID NO: 5519:
US-09-107-532A-5519

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
Db 63 GKVLGL 68

RESULT 889
US-09-097-889-25
; Sequence 25, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-097-889-25

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
| | | | |
Db 288 LGGVLA 293

RESULT 890
US-09-098-079-25
; Sequence 25, Application US/09098079
; Patent No. 6489095
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Clevenger, William
; APPLICANT: Fahy, Bojn P.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON QUANTIFICATION OF
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098,079
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.416
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 380 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
US-09-098-079-25

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 9.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
| | | | |
Db 288 LGGVLA 293

RESULT 891
US-09-902-540-10126
; Sequence 10126, Application US/09902540

; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10126
; LENGTH: 384
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-10126

Query Match 5.1%; Score 6; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 124 LAALAA 129

RESULT 892
US-09-270-767-43634
; Sequence 43634, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43634
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43634

Query Match 5.1%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVTT 11
| | | | |
Db 250 DLEVTT 255

RESULT 893
US-09-712-363-148
; Sequence 148, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844

; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-148

Query Match 5.1%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 255 LAALAA 260

RESULT 894
US-09-938-901A-62
; Sequence 62, Application US/09938901A
; Patent No. 6939949
; GENERAL INFORMATION:
; APPLICANT: Kuramitsu, Seiki
; APPLICANT: Yokoyama, Shigeyuki
; TITLE OF INVENTION: No. 6939949el DNA Repair Enzymes, Nucleic Acids Encoding DNA Repa
; FILE REFERENCE: 11283-013001/PH1261 US
; CURRENT APPLICATION NUMBER: US/09/938,901A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: JP 47762/2001
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 62
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-938-901A-62

Query Match 5.1%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 255 LAALAA 260

RESULT 895
US-09-328-352-6823
; Sequence 6823, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 6823
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6823

Query Match 5.1%; Score 6; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EQAQVI 78
|||
Db 20 EQAQVI 25

RESULT 896
US-09-902-540-16621
; Sequence 16621, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 16621
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-16621

Query Match 5.1%; Score 6; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 93 LAALAA 98

RESULT 897
US-09-489-039A-8995
; Sequence 8995, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8995
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8995

Query Match 5.1%; Score 6; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGK 45
|||
Db 109 IELGK 114

RESULT 898
US-09-489-039A-13498
; Sequence 13498, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13498
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13498

Query Match 5.1%; Score 6; DB 2; Length 397;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 ATQQA 98
| | | | |
Db 308 ATQQA 313

RESULT 899
US-09-489-039A-10656
; Sequence 10656, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10656
; LENGTH: 398
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10656

Query Match 5.1%; Score 6; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GHILG 43
| | | | |
Db 68 GHILG 73

RESULT 900
US-08-834-655-4
; Sequence 4, Application US/08834655
; Patent No. 5968809
; GENERAL INFORMATION:
; APPLICANT: KNUTZON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 18

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RAE-VENTER LAW GROUP, P.C.
; STREET: 260 SHERIDAN AVENUE, P.O. BOX 60039
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/834,655
; APPLICATION NUMBER: US/08/834,655
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER, BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.124.00US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 328-4400
; TELEFAX: (650) 328-4477
; TELEX: N/A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 399 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-834-655-4

Query Match 5.1%; Score 6; DB 1; Length 399;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 253 GVLAAL 258

RESULT 901
US-08-834-033A-4
; Sequence 4, Application US/08834033A
; Patent No. 6075183
; GENERAL INFORMATION:
; APPLICANT: KNUTZON, DEBORAH
; APPLICANT: MUKERJI, PRADIP
; APPLICANT: HUANG, YUNG-SHENG
; APPLICANT: THURMOND, JENNIFER
; APPLICANT: CHAUDHARY, SUNITA
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
; OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LIMBACH AND LIMBACH, L.L.P.
; STREET: 2001 FERRY BUILDING
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/834,033A
; FILING DATE: 11-APR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WARD, MICHAEL R.

REGISTRATION NUMBER: 38,651
REFERENCE/DOCKET NUMBER: CGAB-300.USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-834-033A-4

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 253 GVLAAL 258

RESULT 902
US-09-363-574-4
Sequence 4, Application US/09363574
Patent No. 6136574
GENERAL INFORMATION:
APPLICANT: KNUTZON, DEBORAH
APPLICANT: MURKERJI, PRADIP
APPLICANT: HUANG, YUNG-SHENG
APPLICANT: THURMOND, JENNIFER
APPLICANT: CHAUDHARY, SUNITA
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS
OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS

NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESS: LIMBACH AND LIMBACH L.L.P.
STREET: 2001 FERRY BUILDING
CITY: SAN FRANCISCO
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/363,574
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: WARD, MICHAEL R.
REGISTRATION NUMBER: 38,651
REFERENCE/DOCKET NUMBER: CGAB-202 USA
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 433-4150
TELEFAX: (415) 433-8716
TELEX: N/A

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 399 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-363-574-4

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 253 GVLAAL 258

RESULT 903

US-09-363-526-4
Sequence 4, Application US/09363526
Patent No. 6410288

GENERAL INFORMATION:

APPLICANT: KNUTZON, DEBORAH

APPLICANT: MURKERJI, PRADIP

APPLICANT: HUANG, YUNG-SHENG

APPLICANT: THURMOND, JENNIFER

APPLICANT: CHAUDHARY, SUNITA

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SYNTHESIS

OF LONG CHAIN POLY-UNSATURATED FATTY ACIDS IN PLANTS

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESS: LIMBACH AND LIMBACH L.L.P.

STREET: 2001 FERRY BUILDING

CITY: SAN FRANCISCO

STATE: CA

COUNTRY: USA

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/363,526

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: WARD, MICHAEL R.

REGISTRATION NUMBER: 38,651

REFERENCE/DOCKET NUMBER: CGAB-201 USA

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 433-4150

TELEFAX: (415) 433-8716

TELEX: N/A

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 399 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-363-526-4

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 253 GVLAAL 258

RESULT 904

US-09-330-235-20

Sequence 20, Application US/09330235

Patent No. 6459018

GENERAL INFORMATION:

APPLICANT: Knutzon, Debbie

TITLE OF INVENTION: POLYUNSATURATED FATTY ACIDS IN PLANTS

FILE REFERENCE: MOCO.156.00US

CURRENT APPLICATION NUMBER: US/09/330,235

CURRENT FILING DATE: 1999-06-10

PRIOR APPLICATION NUMBER: 60/089,043

PRIOR FILING DATE: 1998-06-12

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Mortierella alpina
US-09-330-235-20

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 253 GVLAAL 258
|||||

RESULT 905

US-08-403-852D-19
; Sequence 19, Application US/08403852D
; Patent No. 5891695
; GENERAL INFORMATION:

; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Lauret
; APPLICANT: De Crecy-Legard, Valerie

; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner

; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.

; COUNTRY: USA
; ZIP: 20005-3315

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/403,852D
; FILING DATE: 10-MAY-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 93/00923

; FILING DATE: 25-SEP-1993
; APPLICATION NUMBER: FR 92/11441

; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:

; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146

; REFERENCE/DOCKET NUMBER: 03906.0054-00000
; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids

; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
US-08-403-852D-19

Query Match 5.1%; Score 6; DB 1; Length 402;
Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALA 25
Db 222 VLAALA 227
|||||

RESULT 906

US-08-510-646B-20
; Sequence 20, Application US/08510646B
; Patent No. 6077699
; GENERAL INFORMATION:

; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Lauret
; APPLICANT: De Crecy-Legard, Valerie

; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner

; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.

; COUNTRY: USA
; ZIP: 20005-3315

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/510,646B
; FILING DATE: 03-AUG-1995

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/403,852
; FILING DATE: 10-MAY-1995

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 93/00923

; FILING DATE: 25-SEP-1993
; APPLICATION NUMBER: FR 92/11441

; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:

; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146

; REFERENCE/DOCKET NUMBER: 03806.0054-01000
; TELEPHONE: (202) 408-4000

; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids

; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
US-08-510-646B-20

Query Match 5.1%; Score 6; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALA 25
Db 222 VLAALA 227
|||||

RESULT 907

US-09-231-818-19
; Sequence 19, Application US/09231818
; Patent No. 6171846
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; Biosynthesis Of Streptogramins, Nucleotide Sequences
; Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/231,818
; FILING DATE: 25-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-09-231-818-19

Query Match 5.1%; Score 6; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 1e+03; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALA 25
Db 222 VLAALA 227
RESULT 908
US-09-635-359B-19
; Sequence 19, Application US/09635359B
; Patent No. 6670157
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanche, Francis

Crouzet, Joel
Jacques, Nathalie
Lacroix, Patricia
Thibaut, Denis
Zagorec, Monique
Debussche, Laurent
De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; Biosynthesis Of Streptogramins, Nucleotide Sequences
; Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/635,359B
; FILING DATE: 09-Aug-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/231,818
; FILING DATE: 15-JAN-1999
; APPLICATION NUMBER: US 08/403,852
; FILING DATE: 10-MAY-1995
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-635-359B-19

Query Match 5.1%; Score 6; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 1e+03; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALA 25
Db 222 VLAALA 227
RESULT 909
US-09-438-185A-567
; Sequence 567, Application US/09438185A
; Patent No. 6822071
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-000411US
; CURRENT APPLICATION NUMBER: US/09/438,185A

; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 567
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: Cpn0565
US-09-438-185A-567

Query Match 5.1%; Score 6; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 77 LGGVLA 82
|||||

RESULT 910

US-09-328-352-5318
; Sequence 5318, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5318
; LENGTH: 407
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5318

Query Match 5.1%; Score 6; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 1.e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 21
Db 362 LGGVLA 367
|||||

RESULT 911

US-09-712-363-268
; Sequence 268, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593

; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 268
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-268

Query Match 5.1%; Score 6; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
Db 238 GVLAAAL 243
|||||

RESULT 912

US-09-902-540-13562
; Sequence 13562, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYCOCCUS XANTHUS Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13562
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13562

Query Match 5.1%; Score 6; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TWVLLG 18
Db 215 TWVLLG 220
|||||

RESULT 913

US-09-489-039A-10457
; Sequence 10457, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709-2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10457
; LENGTH: 415

; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10457

Query Match 5.1%; Score 6; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 390 LLGGVL 395

RESULT 914

US-09-740-288A-28
; Sequence 28, Application US/09740288A
; Patent No. 6849783
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; APPLICANT: Kinney, Anthony
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Orozco, Emil
; FILE OF INVENTION: PLANT BIOTIN SYNTHASE
; FILE REFERENCE: BB1429 US NA
; CURRENT APPLICATION NUMBER: US/09/740,288A
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/172929
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 28
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Glycine max
US-09-740-288A-28

Query Match 5.1%; Score 6; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
DB 386 KVLGLL 391

RESULT 915

US-08-462-467B-6
; Sequence 6, Application US/08462467B
; Patent No. 6210899
; GENERAL INFORMATION:
; APPLICANT: Rosenbaum, Jan S
; TITLE OF INVENTION: The Use of a BMP Protein Receptor
; TITLE OF INVENTION: Complex for Screening Bone Metabolism Actives and Cells
; TITLE OF INVENTION: Co-Transfected With a Type II BMP Receptor and a Type I
; TITLE OF INVENTION: BMP Receptor
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESS: The Procter & Gamble Company
; STREET: 11810 East Miami River Road
; CITY: Ross
; STATE: OH
; COUNTRY: USA
; ZIP: 45061
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,467B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Hersko, Bart S.
; REGISTRATION NUMBER: 32,572
; REFERENCE/DOCKET NUMBER: 5474R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 627-0633
; TELEFAX: (513) 627-0260
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 417 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
US-08-462-467B-6

Query Match 5.1%; Score 6; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 64 LAALAA 69

RESULT 916

US-09-543-681A-7435
; Sequence 7435, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7435
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-7435

Query Match 5.1%; Score 6; DB 2; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GHIELG 43
DB 91 GHIELG 96

RESULT 917

US-09-614-912-196
; Sequence 196, Application US/09614912
; Patent No. 6677502
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Pamodu, Omolayo O.
; APPLICANT: Lee, Jian Ming
; APPLICANT: Sakai, Hajime
; APPLICANT: Weng, Zude
; APPLICANT: Caimi, Perry G
; APPLICANT: Anderson, Shawn
; TITLE OF INVENTION: Plant Metabolism Genes
; FILE REFERENCE: BB1378 US NA
; CURRENT APPLICATION NUMBER: US/09/614,912
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: 60/143,401

;; PRIOR FILING DATE: 1999-07-12
;; PRIOR APPLICATION NUMBER: 60/143,412
;; PRIOR FILING DATE: 1999-07-12
;; PRIOR APPLICATION NUMBER: 60/146,650
;; PRIOR FILING DATE: 1999-07-30
;; PRIOR APPLICATION NUMBER: 60/170,906
;; PRIOR FILING DATE: 1999-12-15
;; PRIOR APPLICATION NUMBER: 60/172,959
;; PRIOR FILING DATE: 1999-12-21
;; PRIOR APPLICATION NUMBER: 60/172,946
;; PRIOR FILING DATE: 1999-12-21
;; NUMBER OF SEQ ID NOS: 204
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 196
;; LENGTH: 418
;; TYPE: PRT
;; ORGANISM: Glycine max
US-09-614-912-196

Query Match 5.1%; Score 6; DB 2; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 54 VLYQQY 59
|||
Db 129 VLYQQY 134

RESULT 918
US-08-065-844A-2
; Sequence 2, Application US/08065844A
; Patent No. 6333168
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M.
; APPLICANT: Basler, Konrad
; APPLICANT: Yoneda, Toshiya
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/065,844A
; FILING DATE: 19930520
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0576/40314
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 427 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-065-844A-2

Query Match 5.1%; Score 6; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 19 GVLAAL 24
|||
Db 5 GVLAAL 10
RESULT 919
US-10-002-278-2
; Sequence 2, Application US/10002278
; Patent No. 6916913
; GENERAL INFORMATION:
; APPLICANT: Jessell, Thomas M.
; APPLICANT: Basler, Konrad
; APPLICANT: Yamada, Toshiya
; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF DORSALIN-1
; FILE REFERENCE: 0575/40314-A
; CURRENT APPLICATION NUMBER: US/10/002,278
; CURRENT FILING DATE: 2001-11-02
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Chick
US-10-002-278-2

Query Match 5.1%; Score 6; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||
Db 5 GVLAAL 10

RESULT 920
US-09-252-991A-16841
; Sequence 16841, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16841
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16841

Query Match 5.1%; Score 6; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 84 GKVLGL 89
|||
Db 337 GKVLGL 342

RESULT 921
US-09-134-000C-4819
; Sequence 4819, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; PRIOR FILING DATE: 1998-08-13
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4819
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Enterococcus faecalis

US-09-134-000C-4819

Query Match 5.1%; Score 6; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
DB 393 GGVLA 398

RESULT 922
US-09-489-039A-13901
; Sequence 13901, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13901
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13901

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
DB 393 GGVLA 398

RESULT 923
US-09-489-039A-13901
; Sequence 13901, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13901
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13901

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 140 LAALAA 145

RESULT 924
US-09-902-540-10023
; Sequence 10023, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS GENOME SEQUENCES AND USES THEREOF
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10023
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

US-09-902-540-10023

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 140 LAALAA 145

RESULT 925
US-09-605-703B-1244
; Sequence 1244, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1244
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

US-09-605-703B-1244

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 926
US-09-902-540-10023
; Sequence 10023, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS GENOME SEQUENCES AND USES THEREOF
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10023
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

US-09-902-540-10023

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 927
US-09-605-703B-1244
; Sequence 1244, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1244
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

US-09-605-703B-1244

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-902-540-10023

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 AQVIAH 80
DB 344 AQVIAH 349

RESULT 924
US-09-902-540-15774
; Sequence 15774, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS GENOME SEQUENCES AND USES THEREOF
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15774
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

US-09-902-540-15774

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 925
US-09-605-703B-1244
; Sequence 1244, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1244
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

US-09-605-703B-1244

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 926
US-09-902-540-15774
; Sequence 15774, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS GENOME SEQUENCES AND USES THEREOF
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15774
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

US-09-902-540-15774

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 927
US-09-605-703B-1244
; Sequence 1244, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1244
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

US-09-605-703B-1244

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 928
US-09-902-540-15774
; Sequence 15774, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: MYXOCOCCUS XANTHUS GENOME SEQUENCES AND USES THEREOF
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15774
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

US-09-902-540-15774

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 LGGKPA 47
DB 350 LGGKPA 355

RESULT 929
US-09-605-703B-1244
; Sequence 1244, Application US/09605703B
; Patent No. 6962989
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL
; FILE REFERENCE: BGI-129CP
; CURRENT APPLICATION NUMBER: US/09/605,703B
; CURRENT FILING DATE: 2000-06-27
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-07-08
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 2934
; SEQ ID NO 1244
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum

US-09-605-703B-1244

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 19 GVLAAL 24
|||||
Db 86 GVLAAL 91

RESULT 926

US-09-079-030-220
; Sequence 220, Application US/09079030
; Patent No. 6635623
; GENERAL INFORMATION:
; APPLICANT: Guevera, Jr., Juan G.
; APPLICANT: Hoogveen, Ron C.
; APPLICANT: Moore, Paul J.
; TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY
; TITLE OF INVENTION: VECTORS FOR TRANSFECTION OF EUKARYOTIC CELLS
; NUMBER OF SEQUENCES: 229
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/079,030
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: ARAG:003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 220:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

US-09-079-030-220

Query Match 5.1%; Score 6; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLL 90
|||||
Db 203 KVLGLL 208

RESULT 927

US-09-252-991A-22522
; Sequence 22522, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22522
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22522

Query Match 5.1%; Score 6; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
|||||
Db 218 VLAALA 223

RESULT 928

US-09-252-991A-24390
; Sequence 24390, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24390
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24390

Query Match 5.1%; Score 6; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 54 LAALAA 59

RESULT 929

US-09-902-540-15543
; Sequence 15543, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 15543
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Myxococcus xanthus

US-09-902-540-15543

Query Match 5.1%; Score 6; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

;; TITLE OF INVENTION: POLYMERASE ENHANCING FACTOR (PEF)

```

, PRIOR APPLICATION DATA:
, APPLICATION NUMBER: 08/822,774
, FILING DATE: <Unknown>
, ATTORNEY/AGENT INFORMATION:
, NAME: KULIK, David J.
, REGISTRATION NUMBER: 36,576
, REFERENCE/DOCKET NUMBER: 1486/43163
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: (202) 628-8800
, TELEFAX: (202) 628-8844
, INFORMATION FOR SEQ ID NO: 52:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 444 amino acids
, TYPE: amino acid
, TOPOLOGY: unknown
, MOLECULE TYPE: protein
, HYPOTHETICAL: NO
, ANTI-SENSE: NO
, ORIGINAL SOURCE:
, ORGANISM: Escherichia coli
, SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-632-711-52

```

Query Match 5.1%; Score 6; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels

Qy 38 GHIELG 43
Db 98 GHIELG 103

RESULT 934

US-09-632-703B-52
; Sequence 52, Application US/09632703B
; Patent No. 6379553
; GENERAL INFORMATION:
; APPLICANT: HOGREFE, Holly
; TITLE OF INVENTION: Polymerase Enhancing Factor (PEF) Extracts, PEF
; Protein Complexes, Isolated PEF Proteins, and Methods for
; Identifying Same
; NUMBER OF SEQUENCES: 61
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flinsagan, Henderson, Farabow, Garrett, & Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: <Unknown>
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/632,703B
; FILING DATE: 24-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/822,774
; FILING DATE: 21-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: BARKER, M. Paul
; REGISTRATION NUMBER: 32,013
; REFERENCE/DOCKET NUMBER: 4121.0116-04
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 849-6613
; TELEFAX: (650) 849-6666
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 444 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein

? HYPOTHETICAL: NO
 ? ANTI-SENSE: NO
 ? ORIGINAL SOURCE:
 ? ORGANISM: Escherichia coli
 ? SEQUENCE DESCRIPTION: SEQ ID NO: 52:
 US-09-632-7038-52

Query Match 5.1%; Score 6; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels

Qy 38 GHIELG 43
p_b 98 GHIELG 103

RESULT 935

US-09-632-702-52
; Sequence 52, Application US/09632702
; Patent No. 6444428

GENERAL INFORMATION:
APPLICANT: HOGREFE, Holly
TITLE OF INVENTION: Polymerase Enhancing Factor (PEF)
Extracts, p37 Protein Complexes, Isolated PEF Proteins,
and Methods for Purifying and Identifying Same
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: David J. Kulik, Evenson, McKeown, Edwards &
Lenahan, P.L.L.C.
STREET: 1200 G Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005

ZIP: 20005
COMPUTER READABLE FORM.

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/632,702
  FILING DATE: 04-Aug-2000
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US/08/822,774
  FILING DATE: 21-MAR-1997
ATTORNEY/AGENT INFORMATION:
  NAME: KULIK, David J.
  REGISTRATION NUMBER: 36,576
  REFERENCE/DOCKET NUMBER: 1486/43163
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (202) 628-8800
  TELEFAX: (202) 628-8844
INFORMATION FOR SEQ ID NO: 52:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 444 amino acids
    TYPE: amino acid
    TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
  ORGANISM: Escherichia coli
  SEQUENCE DESCRIPTION: SEQ ID NO: 52:
US-09-632-702-52

```

Query Match 5.1%; Score 6; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels

QY		38 GHIELG 43
Db		98 GHIELG 103

RESULT 936
US-09-252-991A-20496
; Sequence 20496, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20496
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (31)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-20496

Query Match 5.1%; Score 6; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
|||
Db 394 VLLGGV 399

RESULT 937
US-09-399-003-52
; Sequence 52, Application US/09399003
; Patent No. 6734293
; GENERAL INFORMATION:
; APPLICANT: Hogrefe, Holly
; APPLICANT: Hansen, Connie J
; TITLE OF INVENTION: Polymerase Enhancing Factor (PEF) Extracts, PEF Protein Complexes
; FILE REFERENCE: 4121.0116-02
; CURRENT APPLICATION NUMBER: US/09/399,003
; CURRENT FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: PCT/US98/05497
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: US 08/957,709
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: US 08/822,774
; PRIOR FILING DATE: 1997-03-21
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Pyrococcus furiosus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(444)
; OTHER INFORMATION: "X" represents any amino acid
US-09-399-003-52

Query Match 5.1%; Score 6; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GHIELG 43
|||
Db 98 GHIELG 103

RESULT 938
US-09-047-026A-6
; Sequence 6, Application US/09047026A
; Patent No. 5989897
; GENERAL INFORMATION:
; APPLICANT: Pillus, Lorraine
; APPLICANT: Clarke, Astrid
; APPLICANT: Lowell, Joanna
; APPLICANT: Jacobson, Sandra
; APPLICANT: Reifsnnyder, Cheryl
; TITLE OF INVENTION: Yeast Silencing Genes, Proteins and
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/047,026A
; FILING DATE: 24-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/042,375
; FILING DATE: 24-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 1-97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-047-026A-6

Query Match 5.1%; Score 6; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.1e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 DKEVLY 56
|||
Db 75 DKEVLY 80

RESULT 939
US-09-543-681A-5700
; Sequence 5700, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5700

```
; LENGTH: 446
; TYPE: ERT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5700

Query Match          5.1%; Score 6; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
Db 12 LLGGVL 17

RESULT 940
US-09-902-540-12215
; Sequence 12215, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12215
; LENGTH: 446
; TYPE: ERT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12215

Query Match          5.1%; Score 6; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
Db 163 LLGGVL 168

RESULT 941
US-09-071-035-256
; Sequence 256, Application US/09071035
; Patent No. 6448043
; GENERAL INFORMATION:
; APPLICANT: Gil H. Choi
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 496
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/071,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

Query Match          5.1%; Score 6; DB 2; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Gaps 0;
```

```
; NAME: A. Anders Brookes
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB369P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 256:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-071-035-256

Query Match          5.1%; Score 6; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 APYIEQ 74
Db 67 APYIEQ 72

RESULT 942
US-10-206-576-256
; Sequence 256, Application US/10206576
; Patent No. 6913907
; GENERAL INFORMATION:
; APPLICANT: Choi et al.
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 497
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-R
; COMPUTER: Dell Latitude
; OPERATING SYSTEM: Windows 98
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,576
; FILING DATE: 29-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/071,035
; FILING DATE: 1998-05-04
; APPLICATION NUMBER: US 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: US 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: US 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Hyman, Mark J.
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB369P1D1
; INFORMATION FOR SEQ ID NO: 256:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 256:
US-10-206-576-256

Query Match          5.1%; Score 6; DB 2; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
```



```
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 69 APYIEQ 74
Db 67 APYIEQ 72
|||||

RESULT 943
US-09-248-796A-20296
; Sequence 20296, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20296
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20296

Query Match 5.1%; Score 6; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 25 AAYCLS 30
Db 142 AAYCLS 147
|||||

RESULT 944
US-09-949-016-10146
; Sequence 10146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10146
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10146

Query Match 5.1%; Score 6; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 LGGVYL 21
Db 264 LGGVYL 269
|||||

RESULT 945
```

```
US-09-252-991A-24836
; Sequence 24836, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24836
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24836

Query Match 5.1%; Score 6; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 41 ELGKGP 46
Db 367 ELGKGP 372
|||||

RESULT 946
US-09-107-532A-6473
; Sequence 6473, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6473:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 455 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
```


US-09-902-540-10552
; Sequence 10552, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902,540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 10552

; LENGTH: 465

; TYPE: PRT

; ORGANISM: Myxococcus xanthus

US-09-902-540-10552

Query Match 5.1%; Score 6; DB 2; Length 465;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

Db 163 VLAALA 168

|||||

RESULT 952

US-09-252-991A-30829

; Sequence 30829, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30829

; LENGTH: 466

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30829

Query Match 5.1%; Score 6; DB 2; Length 466;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20

Db 122 VLLGGV 127

|||||

RESULT 953

US-09-543-681A-7068

; Sequence 7068, Application US/09543681A

; Patent No. 6605709

; GENERAL INFORMATION:

; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/09/543,681A

; CURRENT FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: US 60/128,706

; PRIOR FILING DATE: 1999-04-09

; NUMBER OF SEQ ID NOS: 8344

; SEQ ID NO 7068

; LENGTH: 469

; TYPE: PRT

; ORGANISM: Proteus mirabilis

US-09-543-681A-7068

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 469;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGILL 90

Db 395 KVLGILL 400

|||||

RESULT 954

US-10-033-109-14

; Sequence 14, Application US/10033109

; Patent No. 6833492

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen M.

; APPLICANT: Rafalski, J. Antoni

; APPLICANT: Sakai, Hajime

; TITLE OF INVENTION: Nitrogen Transport Metabolism

; FILE REFERENCE: BB-1210

; CURRENT APPLICATION NUMBER: US/10/033,109

; CURRENT FILING DATE: 2001-12-28

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/384,625

; PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-27

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/098,248

; PRIOR FILING DATE: EARLIER FILING DATE: 28 August 1998

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 14

; LENGTH: 470

; TYPE: PRT

; ORGANISM: Triticum aestivum

US-10-033-109-14

Query Match 5.1%; Score 6; DB 2; Length 470;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGV 21

Db 361 LLGGV 366

|||||

RESULT 955

US-09-489-039A-13479

; Sequence 13479, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 2709.2004001

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 13479

; LENGTH: 472

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13479

Query Match 5.1%; Score 6; DB 2; Length 472;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      17 LGGVLA 22
Db      350 LGGVLA 355

RESULT 956
US-09-902-540-14925
; Sequence 14925, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14925
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14925

Query Match      5.1%; Score 6; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 AALAA 27
Db      34 AALAA 39

RESULT 957
US-09-248-796A-16084
; Sequence 16084, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16084
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Candida albicans
; NAME/KEY: UNSURE
; LOCATION: (50),(429),(449)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-16084

Query Match      5.1%; Score 6; DB 2; Length 473;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      83 KGVLA 88
Db      231 KGVLA 236

RESULT 958
```

```
US-09-252-991A-19246
; Sequence 19246, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19246
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19246

Query Match      5.1%; Score 6; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 22
Db      257 LGGVLA 262

RESULT 959
US-09-270-767-45923
; Sequence 45923, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45923
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-45923

Query Match      5.1%; Score 6; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 LAALAA 26
Db      65 LAALAA 70

RESULT 960
US-09-902-540-9804
; Sequence 9804, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
```

```
; SEQ ID NO 9804
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-9804

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 483;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 94 LAALAA 99

RESULT 961
US-09-270-767-57297
; Sequence 57297, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57297
; LENGTH: 487
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57297

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 487;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
| | | | |
Db 225 GGVLA 230

RESULT 962
US-09-489-039A-10133
; Sequence 10133, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10133
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10133

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 488;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 QAQVIA 79
| | | | |
Db 410 QAQVIA 415

RESULT 963
US-09-252-991A-26250
```

```
; Sequence 26250, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26250
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26250

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 490;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQR 92
| | | | |
Db 154 LGLLQR 159

RESULT 964
US-09-543-681A-6264
; Sequence 6264, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6264
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6264

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 492;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQR 92
| | | | |
Db 94 LGLLQR 99

RESULT 965
US-09-252-991A-23214
; Sequence 23214, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
```

RESULT 968
US-09-489-039A-10027
; Sequence 10027, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10027
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10027

Query Match 5.1%; Score 6; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
| | | | |
Db 69 VLLGGV 74

RESULT 966
US-09-807-258-2
; Sequence 2, Application US/09807258
; Patent No. 6670166
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Arthropod Protein Disulfide Isomerases
; FILE REFERENCE: BB-1253 PCT
; CURRENT APPLICATION NUMBER: US/09/807,258
; CURRENT FILING DATE: 2001-06-11
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 2
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Argiope sp.
US-09-807-258-2

Query Match 5.1%; Score 6; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PKGKVL 87
| | | | |
Db 278 PKGKVL 283

RESULT 967
US-09-807-258-4
; Sequence 4, Application US/09807258
; Patent No. 6670166
; GENERAL INFORMATION:
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Arthropod Protein Disulfide Isomerases
; FILE REFERENCE: BB-1253 PCT
; CURRENT APPLICATION NUMBER: US/09/807,258
; CURRENT FILING DATE: 2001-06-11
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 4
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Scolopendra canidens DS
US-09-807-258-4

Query Match 5.1%; Score 6; DB 2; Length 496;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PKGKVL 87
| | | | |
Db 278 PKGKVL 283

Query Match 5.1%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQR 92
| | | | |
Db 387 LGLLQR 392

RESULT 969
US-09-328-352-8183
; Sequence 8183, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8183
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8183

Query Match 5.1%; Score 6; DB 2; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 WVIVGH 39
| | | | |
Db 350 WVIVGH 355

RESULT 970
US-09-949-016-8062
; Sequence 8062, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
US-09-949-016-8062

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8062
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8062

Query Match 5.1%; Score 6; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
DB 282 LAALAA 287

RESULT 971

US-09-949-016-6538
; Sequence 6538, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6538
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6538

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 505;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||
DB 85 LGGVLA 90

RESULT 972

US-09-252-991A-19050
; Sequence 19050, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19050
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19050

Query Match 5.1%; Score 6; DB 2; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
|||
DB 204 GGVLA 209

RESULT 973

US-09-252-991A-29467
; Sequence 29467, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29467
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29467

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 510;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||
DB 364 LLGGVL 369

RESULT 974

US-09-712-363-227
; Sequence 227, Application US/09712363
; Patent No. 6892139
; GENERAL INFORMATION:
; APPLICANT: Eisenberg, David
; APPLICANT: Rotstein, Sergio H.
; APPLICANT: Marcotte, Edward M.
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
; FILE REFERENCE: 07419-032001
; CURRENT APPLICATION NUMBER: US/09/712,363
; PRIOR FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: PCT/US00/02246
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/179,531
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 60/117,844
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: 60/118,206,
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: 60/126,593
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/134,093
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/134,092
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/165,124
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/165,086
; PRIOR FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 292

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 227
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-712-363-227

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 510;
; Sequence 56, Application US/09833745
; Patent No. 6939541
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 109 VLAALA 114
|||||

RESULT 975
US-09-902-540-13924
; Sequence 13924, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wisegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13924
; LENGTH: 512
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13924

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 512;
; Sequence 30773, Application US/09252991A
; Patent No. 6551795
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 247 LGGVLA 252
|||||

RESULT 976
US-09-252-991A-30773
; Sequence 30773, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30773
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30773

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 513;
; Sequence 60, Application US/09833745
; Patent No. 6939541
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 91 QRATQQ 96
Db 484 QRATQQ 489
|||||

RESULT 977
US-09-833-745-56
; Sequence 56, Application US/09833745
; Patent No. 6939541
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-745-56

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 513;
; Sequence 59, Application US/09833745
; Patent No. 6939541
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 AAPYIE 73
Db 507 AAPYIE 512
|||||

RESULT 978
US-09-833-745-59
; Sequence 59, Application US/09833745
; Patent No. 6939541
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 59
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-833-745-59

Query Match
Best Local Similarity 5.1%; Score 6; DB 2; Length 513;
; Sequence 60, Application US/09833745
; Patent No. 6939541
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 AAPYIE 73
Db 507 AAPYIE 512
|||||

RESULT 979
US-09-833-745-60
; Sequence 60, Application US/09833745
; Patent No. 6939541
```



```
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; TITLE OF INVENTION: BIOACTIVE HISTIDINE AMMONIA LYASE
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-833-745-60

Query Match          5.1%; Score 6; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      68 AAPYIE 73
Db      507 AAPYIE 512

RESULT 980
US-09-833-745-61
; Sequence 61, Application US/09833745
; Patent No. 6939541
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; TITLE OF INVENTION: BIOACTIVE HISTIDINE AMMONIA LYASE
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-833-745-61

Query Match          5.1%; Score 6; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      68 AAPYIE 73
Db      507 AAPYIE 512

RESULT 981
US-09-252-991A-20224
; Sequence 2024, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8809
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8809
```

```
; GENERAL INFORMATION:
; APPLICANT: ROBERTS, JOSEPH
; APPLICANT: SETHURAMAN, NATARAJAN
; APPLICANT: MACALLISTER, THOMAS
; TITLE OF INVENTION: CLONING, OVEREXPRESSION AND THERAPEUTIC USE OF
; TITLE OF INVENTION: BIOACTIVE HISTIDINE AMMONIA LYASE
; FILE REFERENCE: 078728/0106
; CURRENT APPLICATION NUMBER: US/09/833,745
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,770
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-833-745-60

Query Match          5.1%; Score 6; DB 2; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      68 GLLQRA 93
Db      84 GLLQRA 89

RESULT 982
US-09-252-991A-28903
; Sequence 28903, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28903
; LENGTH: 520
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28903

Query Match          5.1%; Score 6; DB 2; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GVLAAL 24
Db      92 GVLAAL 97

RESULT 983
US-09-949-016-8809
; Sequence 8809, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8809
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8809
```

Query Match 5.1%; Score 6; DB 2; Length 521;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 101 LGGVLA 106
|||||

RESULT 984
US-09-252-991A-18092
; Sequence 18092, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18092
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18092

Query Match 5.1%; Score 6; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 436 LAALAA 441
|||||

RESULT 985
US-09-242-913B-15
; Sequence 15, Application US/09242913B
; Patent No. 6551811
; GENERAL INFORMATION:
; APPLICANT: FONTAINE, THIERRY
; APPLICANT: HARTLAND, ROBERT
; APPLICANT: MOUINA, ISABELLE
; APPLICANT: LATGE, JEAN-PAUL
; TITLE OF INVENTION: METHOD FOR SORTING ANTIFUNGAL MOLECULES ACTING ON THE
; TITLE OF INVENTION: GLUCANOSYLTRANSFERASE ACTIVITY
; FILE REFERENCE: 05986-0007
; CURRENT APPLICATION NUMBER: US/09/242,913B
; CURRENT FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: PCT/FR97/01540
; PRIOR FILING DATE: 1997-08-29
; PRIOR APPLICATION NUMBER: 60/024,910
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-242-913B-15

Query Match 5.1%; Score 6; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||

Db 9 LAALAA 14

RESULT 986
US-09-252-991A-26641
; Sequence 26641, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26641
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26641

Query Match 5.1%; Score 6; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
Db 391 GVLAAAL 396
|||||

RESULT 987
US-08-975-762-73
; Sequence 73, Application US/08975762
; Patent No. 6207169
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,762
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.439
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-975-762-73

Query Match 5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||||
Db 202 VLGLLQ 207

RESULT 988

US-09-295-028-73
; Sequence 73, Application US/09295028
; Patent No. 6277381

GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF EHRlichia INFECTION
; FILE REFERENCE: 210121.439C4
; CURRENT APPLICATION NUMBER: US/09/295,028
; CURRENT FILING DATE: 1999-04-20
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Ehrlichia sp.
US-09-295-028-73

Query Match 5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||||
Db 202 VLGLLQ 207

RESULT 989

US-09-106-582-73
; Sequence 73, Application US/09106582
; Patent No. 6306402

GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/106,582
; FILING DATE: 29-JUN-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.439C2

THERAPY

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-106-582-73

Query Match 5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||||
Db 202 VLGLLQ 207

RESULT 990

US-09-159-469-73
; Sequence 73, Application US/09159469
; Patent No. 6607728

GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; TITLE OF INVENTION: THERAPY OF EHRlichia INFECTION
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/159,469
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/106,582
; FILING DATE: 29-JUN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.439C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4900
; TELEFAX: 206-682-6031
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-159-469-73

Query Match 5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||||

```
Db      202 VIGLLQ 207

RESULT 991
US-09-693-542-73
; Sequence 73, Application US/09693542
; Patent No. 6673156
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; TITLE OF INVENTION: AND TREATMENT OF EHRlichia INFECTION
; FILE REFERENCE: 210121.439C6
; CURRENT APPLICATION NUMBER: US/09/693,542
; CURRENT FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 73
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Ehrlichia sp.
US-09-693-542-73

Query Match      5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      86 VIGLLQ 91
      |||||
Db      202 VIGLLQ 207

RESULT 992
US-09-444-711A-4
; Sequence 4, Application US/09444711A
; Patent No. 6764833
; GENERAL INFORMATION:
; APPLICANT: Yeatman, Timothy J.
; APPLICANT: Irby, Rosalyn B.
; TITLE OF INVENTION: Mutated SRC Oncogene Composition and Methods
; FILE REFERENCE: USF-T136
; CURRENT APPLICATION NUMBER: US/09/444,711A
; CURRENT FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 530
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(530)
; OTHER INFORMATION: amino acid sequence of the mutant c-Src polypeptide encoded
; OTHER INFORMATION: by the mutant c-Src coding region
US-09-444-711A-4

Query Match      5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      26 AYCLSV 31
      |||||
Db      186 AYCLSV 191

RESULT 993
US-08-339-152A-32
; Sequence 32, Application US/08339152A
; Patent No. 5643726
; GENERAL INFORMATION:
; APPLICANT: Tanzi, Rudolph E.
```

```
; APPLICANT: Kovacs, Dora M.
; TITLE OF INVENTION: Methods For Modulating Transcription
; TITLE OF INVENTION: From The Amyloid -Protein Precursor (APP) Promoter
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 New York Ave., NW, Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,152A
; FILING DATE: 10-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 0609.4120000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 532 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-339-152A-32

Query Match      5.1%; Score 6; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 LAALAA 26
      |||||
Db      22 LAALAA 27

RESULT 994
US-07-820-011A-2
; Sequence 2, Application US/07820011A
; Patent No. 5336615
; GENERAL INFORMATION:
; APPLICANT: Bell, Leonard
; APPLICANT: Madri, Joseph A.
; APPLICANT: Warren, Stephen L.
; APPLICANT: Luthringer, Daniel J.
; TITLE OF INVENTION: Genetically Engineered
; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
; TITLE OF INVENTION: Migration
; TITLE OF INVENTION: and Plasminogen Activator Activity
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC XT
; OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,011A
; FILING DATE: 19920106
```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: LB-101
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: AMINO ACID
TOPOLOGY: Linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
FRAGMENT TYPE: Complete Sequence
ORIGINAL SOURCE:
ORGANISM: Gallus, gallus
PUBLICATION INFORMATION:
AUTHORS: Takeya, Tatsuo
TITLE: Structure and Sequence of the
TITLE: Cellular Gene Homologous to the RSV src
TITLE: Gene and the Mechanism for Generating the
TITLE: Transforming Virus
JOURNAL: Cell
VOLUME: 32
PAGES: 881-890
DATE: March, 1983
US-07-820-011A-2

Query Match 5.1%; Score 6; DB 1; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
Db 183 AYCLSV 188

RESULT 995
US-09-470-881-3
Sequence 3, Application US/09470881
Patent No. 6685938
GENERAL INFORMATION:
APPLICANT: CHERESH, David A.
APPLICANT: ELICEIRI, Brian
TITLE OF INVENTION: METHODS AND COMPOSITIONS USEFUL FOR MODULATION OF
TITLE OF INVENTION: ANGIOGENESIS AND VASCULAR PERMEABILITY USING SRC OR
FILE REFERENCE: TSRI 651.2
CURRENT APPLICATION NUMBER: US/09/470,881
CURRENT FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US99/11780
PRIOR FILING DATE: 1999-05-28
PRIOR APPLICATION NUMBER: 60/087,220
PRIOR FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 533
TYPE: PRT
ORGANISM: Chicken
US-09-470-881-3

Query Match 5.1%; Score 6; DB 2; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
Db 183 AYCLSV 188

RESULT 996
PCT-US93-00445-2
Sequence 2, Application PC/TUS9300445
GENERAL INFORMATION:
APPLICANT: Bell, Leonard
APPLICANT: Madri, Joseph A.
APPLICANT: Warren, Stephen L.
APPLICANT: Luthringer, Daniel J.
TITLE OF INVENTION: Genetically Engineered
TITLE OF INVENTION: Endothelial Cells
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 760 Kb storage
COMPUTER: DELL 486/50
OPERATING SYSTEM: DOS 5.0
SOFTWARE: Displaywrite 3
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/00445
FILING DATE: 19930105
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/820,011
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-101PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: AMINO ACID
TOPOLOGY: Linear
MOLECULE TYPE: Protein
HYPOTHETICAL: NO
FRAGMENT TYPE: Complete Sequence
ORIGINAL SOURCE:
ORGANISM: Gallus, gallus
AUTHORS: Takeya, Tatsuo
AUTHORS: Hanafusa, Hidesaburo
TITLE: Structure and Sequence of the
TITLE: Cellular Gene Homologous to the RSV src
TITLE: Gene and the Mechanism for Generating the
TITLE: Transforming Virus
JOURNAL: Cell
VOLUME: 32
PAGES: 881-890
DATE: March, 1983
PCT-US93-00445-2

Query Match 5.1%; Score 6; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
Db 183 AYCLSV 188

RESULT 997
US-09-902-540-13260
Sequence 13260, Application US/09902540

```
; Patent No. 5833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13260
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-13260

Query Match          5.1%; Score 6; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      20 VLAALA 25
Db      456 VLAALA 461

; RESULT 998
; US-07-820-011A-4
; Sequence 4, Application US/07820011A
; Patent No. 5336615
; GENERAL INFORMATION:
; APPLICANT: Bell, Leonard
; APPLICANT: Madril, Joseph A.
; APPLICANT: Warren, Stephen L.
; APPLICANT: Luchringer, Daniel J.
; TITLE OF INVENTION: Genetically Engineered
; TITLE OF INVENTION: Endothelial Cells Exhibiting Enhanced
; TITLE OF INVENTION: Migration
; TITLE OF INVENTION: and Plasminogen Activator Activity
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb storage
; COMPUTER: IBM PC XT
; OPERATING SYSTEM: PC-DOS/MS-DOS 2.10
; SOFTWARE: Displaywrite 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,011A
; FILING DATE: 19920106
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Klee, Maurice M.
; REGISTRATION NUMBER: 30,399
; REFERENCE/DOCKET NUMBER: LB-101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203) 255 1400
; TELEFAX: (203) 254 1101
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 536 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: Linear
; MOLECULE TYPE: Protein
; HYPOTHETICAL: NO
; FRAGMENT TYPE: Complete Sequence
```

```
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; PUBLICATION INFORMATION:
; AUTHORS: Anderson, Stephen K.
; AUTHORS: Gibbs, Carol P.
; AUTHORS: Tanaka, Akio
; AUTHORS: Kung, Hsing-jien
; AUTHORS: Fujita, Donald J.
; TITLE: Human Cellular src Gene:
; TITLE: Nucleotide Sequence and Derived Amino
; TITLE: Acid Sequence of the Region Coding for
; TITLE: the Carboxy-Terminal Two-Thirds of
; TITLE: pp60c-src
; JOURNAL: Molecular and Cellular Biology
; VOLUME: 5
; ISSUE: 5
; PAGES: 1122-1129
; DATE: May, 1985
; PUBLICATION INFORMATION:
; AUTHORS: Tanaka, Akio
; AUTHORS: Gibbs, Carol P.
; AUTHORS: Arthur, Richard R.
; AUTHORS: Anderson, Stephen K.
; AUTHORS: Kung, Hsing-jien
; AUTHORS: Fujita, Donald J.
; TITLE: DNA Sequence Encoding the
; TITLE: Amino-Terminal Region of the Human c-src
; TITLE: Protein: Implications of Sequence
; TITLE: Divergence among src-Type Kinase
; TITLE: Oncogenes
; JOURNAL: Molecular and Cellular Biology
; VOLUME: 7
; ISSUE: 5
; PAGES: 1978-1983
; DATE: May, 1987
; US-07-820-011A-4

Query Match          5.1%; Score 6; DB 1; Length 536;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      26 AYCLSV 31
Db      186 AYCLSV 191

; RESULT 999
; US-08-426-509A-13
; Sequence 13, Application US/08426509A
; Patent No. 6326469
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Gishizsky, Mikhail
; APPLICANT: Sures, Irman G.
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
; TITLE OF INVENTION: TYROSINE KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York,
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,509A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
```

;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/232,545
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Coruzzi, Laura A
;; REGISTRATION NUMBER: 30,742
;; REFERENCE/DOCKET NUMBER: 7683-0074-999
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-790-9090
;; TELEFAX: 212-869-9741
;; TELEX: 66141 PENNIE
;; INFORMATION FOR SEQ ID NO: 13:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 536 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: unknown
;; TOPOLOGY: unknown
;; MOLECULE TYPE: No. 6326469e
US-08-426-509A-13

Query Match 5.1%; Score 6; DB 2; Length 536;
Best Local Similarity 100.0%; Pred.No. 1.4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
Db 186 AYCLSV 191

Search completed: January 27, 2006, 19:29:10
Job time : 40 secs

RESULT 1000
US-08-232-545-13
; Sequence 13, Application US/08232545
; Patent No. 6506578
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Gishizeky, Mikhail
; APPLICANT: Sures, Irman G.
; TITLE OF INVENTION: No. 6506578el Megakaryocytic Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,545
; FILING DATE: 22-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 536 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-232-545-13

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 19:23:06 ; Search time 16 Seconds
(without alignments)
709,599 Million cell updates/sec

Title: US-09-638-693A-36_COPY_16_133

Perfect score: 118

Sequence: 1 ACMSADLEVTSTWLLGGV.....VIEPIVTTNWQKLEAFMKH 118

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 2000 summaries

Database :

PIR 80:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	70.3	209	2	PC1306
2	47	39.8	142	2	PC1307
3	13	11.0	492	2	PS0326
4	13	11.0	716	2	JQ1366
5	13	11.0	3011	1	GNWVC3
6	13	11.0	3011	1	GNWVC3
7	13	11.0	3011	1	S40770
8	12	10.2	3010	1	A45573
9	12	10.2	3010	1	GNWVC3
10	12	10.2	3010	1	GNWVC3
11	12	10.2	3010	1	GNWVC3
12	12	10.2	3010	1	S18030
13	9	7.6	280	2	PC4219
14	9	7.6	876	2	PC2219
15	9	7.6	3014	1	JC5620
16	8	6.8	327	2	S61982
17	8	6.8	342	2	AD3450
18	8	6.8	363	1	E69889
19	8	6.8	399	2	E70598
20	8	6.8	420	2	C84201
21	7	5.9	80	2	S54909
22	7	5.9	96	2	H83065
23	7	5.9	107	2	G83545
24	7	5.9	125	2	S35629
25	7	5.9	128	2	B70647
26	7	5.9	159	2	D95864
27	7	5.9	172	2	B84260
28	7	5.9	171	2	C72737
29	7	5.9	188	2	AE3363

30	7	5.9	203	1	WMBPTB
31	7	5.9	204	2	G95276
32	7	5.9	217	2	B81067
33	7	5.9	217	2	G81801
34	7	5.9	222	2	B95207
35	7	5.9	222	2	B98072
36	7	5.9	228	2	C83663
37	7	5.9	234	2	D96932
38	7	5.9	235	2	E65082
39	7	5.9	245	2	AG2300
40	7	5.9	295	2	JC5671
41	7	5.9	295	2	C71107
42	7	5.9	303	2	AD3473
43	7	5.9	306	2	F72606
44	7	5.9	306	2	F70603
45	7	5.9	359	2	F82541
46	7	5.9	360	2	T00882
47	7	5.9	362	2	T34921
48	7	5.9	366	2	S77203
49	7	5.9	370	2	B72594
50	7	5.9	383	2	F88486
51	7	5.9	389	2	F86268
52	7	5.9	389	2	T14412
53	7	5.9	389	2	G83413
54	7	5.9	390	2	A72698
55	7	5.9	396	2	B75290
56	7	5.9	396	2	AI3554
57	7	5.9	410	2	C84176
58	7	5.9	414	2	D82487
59	7	5.9	417	2	T11387
60	7	5.9	419	2	B70753
61	7	5.9	427	2	D81784
62	7	5.9	428	1	Q4ECAD
63	7	5.9	428	2	D86003
64	7	5.9	428	2	F91157
65	7	5.9	431	2	G81871
66	7	5.9	431	2	A81150
67	7	5.9	450	2	E70590
68	7	5.9	453	2	C82066
69	7	5.9	471	2	C82825
70	7	5.9	491	2	S71862
71	7	5.9	493	2	T34092
72	7	5.9	500	2	AF2325
73	7	5.9	510	2	G72464
74	7	5.9	510	2	A87482
75	7	5.9	515	2	F96829
76	7	5.9	530	2	D82063
77	7	5.9	538	1	Q0ECD7
78	7	5.9	559	2	B84213
79	7	5.9	582	2	G91231
80	7	5.9	582	2	F86078
81	7	5.9	585	2	H83941
82	7	5.9	592	2	H97105
83	7	5.9	603	2	H75272
84	7	5.9	605	2	S74455
85	7	5.9	617	2	F71019
86	7	5.9	645	2	A75390
87	7	5.9	665	2	AU0411
88	7	5.9	729	2	E81410
89	7	5.9	733	2	A95421
90	7	5.9	740	2	F71369
91	7	5.9	797	2	D70581
92	7	5.9	815	2	I57487
93	7	5.9	816	2	S16328
94	7	5.9	818	2	A48858
95	7	5.9	820	2	A40204
96	7	5.9	870	2	F90899
97	7	5.9	881	2	F90997
98	7	5.9	881	2	F83530
99	7	5.9	884	2	D85817
100	7	5.9	914	2	S18942
101	7	5.9	974	1	PXINPD
102	7	5.9	1385	2	S61236

gene 10 protein -
probable transcrip
conserved hypothet
probable lipoprote
aquaporin 2, water
hypothetical prote
ABC transporter, A
hypothetical prote
hypothetical prote
methionyl aminopep
probable methionyl
methyltransferase
probable methanol
hypothetical prote
dihydroxotrate deh
hypothetical prote
probable methyltra
hypothetical prote
hypothetical prote
protein F20H11.5 l
aminoalcoholphosph
ethanolaminephosph
probable MRS trans
hypothetical prote
hypothetical prote
hypothetical prote
xylose transport s
oxalate/formate an
multidrug resistan
NADH2 dehydrogenas
probable efflux pr
probable integral
damX protein (arob
hypothetical prote
probable membrane
probable histidine
histidyl-tRNA synt
3-phosphoshikimate
magnesium transport
UDP-N-acetylmurama
protein disulfide-
hypothetical prote
NADH dehydrogenase
hypothetical prote
conserved hypothet
hypothetical prote
probable thiamin A
tnaB protein - Bsc
hypothetical prote
probable frv opero
probable frv opero
ABC transporter (A
carbon starvation
probable nucleic a
ABC-type transport
hypothetical prote
NADH2 dehydrogenas
ferrichrome transp
transcription regu
probable oxidoredu
conserved hypothet
probable cation-tr
Na+/H+-exchanging
Na+/H+-exchanging
Na+/H+-exchanging
Na+/H+-exchanging
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
H+-exporting ATPas
major capsid prote

103	7	5.9	1392	2	T51947	176	6	5.1	154	2	B96728	hypothetical prote
104	7	5.9	1446	1	A45344	177	6	5.1	155	2	G87445	hypothetical prote
105	7	5.9	1460	1	EBBERF	178	6	5.1	160	2	F83541	probable glutathio
106	7	5.9	2276	2	T00076	179	6	5.1	161	2	S36491	S6 protein - human
107	7	5.9	3033	1	JQ1303	180	6	5.1	162	1	YAZ051	antigen 5.1 precu
108	6	5.1	18	2	S09026	181	6	5.1	162	2	A25780	blood-stage antige
109	6	5.1	20	2	S09025	182	6	5.1	162	2	A26769	antigen 5.1 precu
110	6	5.1	20	2	S09023	183	6	5.1	162	2	H75435	hypothetical prote
111	6	5.1	20	2	S09022	184	6	5.1	163	2	T02438	dehydration-induce
112	6	5.1	22	2	S09021	185	6	5.1	164	2	E89955	hypothetical prote
113	6	5.1	27	2	A45140	186	6	5.1	168	2	G83428	hypothetical prote
114	6	5.1	41	2	P00561	187	6	5.1	170	1	F70660	hypothetical prote
115	6	5.1	41	2	P00560	188	6	5.1	170	2	F43256	hypothetical prote
116	6	5.1	41	2	P00563	189	6	5.1	170	2	F82437	probable rRNA meth
117	6	5.1	41	2	P00564	190	6	5.1	170	2	AE1817	hypothetical prote
118	6	5.1	41	2	P00562	191	6	5.1	174	2	B75013	phosphoribosylamin
119	6	5.1	41	2	P00565	192	6	5.1	177	2	B90988	partial probable s
120	6	5.1	51	1	Q1BP47	193	6	5.1	177	2	E85833	partial probable s
121	6	5.1	61	1	AXLQ2	194	6	5.1	177	2	C95922	hypothetical prote
122	6	5.1	69	2	D91257	195	6	5.1	179	2	H72897	hypothetical prote
123	6	5.1	69	2	H86097	196	6	5.1	180	2	T50598	probable membrane
124	6	5.1	69	2	D65212	197	6	5.1	181	2	C83435	GTP cyclohydrolase
125	6	5.1	71	2	E84364	198	6	5.1	181	2	A86751	prophage p12 prote
126	6	5.1	81	2	G70534	199	6	5.1	182	2	F83175	conserved hypothet
127	6	5.1	81	2	AC1829	200	6	5.1	183	2	D83069	conserved hypothet
128	6	5.1	83	2	B84210	201	6	5.1	185	2	F75255	hypothetical prote
129	6	5.1	83	2	T35070	202	6	5.1	185	2	S75549	hypothetical prote
130	6	5.1	89	2	T44917	203	6	5.1	186	2	D83217	GTP cyclohydrolase
131	6	5.1	91	2	F70630	204	6	5.1	186	2	C95083	conserved hypothet
132	6	5.1	92	2	S41724	205	6	5.1	188	2	S49192	GCR 1 protein - fr
133	6	5.1	94	2	D90249	206	6	5.1	190	2	D82151	hypothetical prote
134	6	5.1	98	2	F83404	207	6	5.1	190	2	AG2668	biotin synthesis B
135	6	5.1	100	2	S44892	208	6	5.1	190	2	S97450	probable biotin sy
136	6	5.1	102	2	T32603	209	6	5.1	191	2	E87519	RnfB/PfpI family p
137	6	5.1	102	2	T31879	210	6	5.1	195	2	F82252	RnfB-related prote
138	6	5.1	102	2	A12693	211	6	5.1	197	2	A83092	probable phosphope
139	6	5.1	104	2	H87548	212	6	5.1	197	2	A30013	plasma retinol-bin
140	6	5.1	106	1	IXL2	213	6	5.1	197	2	F75436	conserved hypothet
141	6	5.1	106	2	A43301	214	6	5.1	197	2	G97950	conserved hypothet
142	6	5.1	106	2	A43301	215	6	5.1	198	2	D87484	glutamine amidotra
143	6	5.1	108	2	C87494	216	6	5.1	198	2	C84237	hypothetical prote
144	6	5.1	108	2	C72392	217	6	5.1	198	2	AB3435	hypothetical prote
145	6	5.1	110	2	A70596	218	6	5.1	199	2	AB0236	probable Na(+)-tra
146	6	5.1	112	2	D75445	219	6	5.1	199	2	T47023	hypothetical prote
147	6	5.1	113	2	I40680	220	6	5.1	199	2	S15164	hypothetical prote
148	6	5.1	115	2	D75507	221	6	5.1	199	2	T36622	hypothetical prote
149	6	5.1	115	2	E95174	222	6	5.1	201	2	T43151	hypothetical prote
150	6	5.1	115	2	G98040	223	6	5.1	203	2	F87540	hypothetical prote
151	6	5.1	117	2	F95862	224	6	5.1	203	2	A99865	hypothetical prote
152	6	5.1	117	2	F89756	225	6	5.1	204	2	H85753	partial probable p
153	6	5.1	119	2	D72520	226	6	5.1	206	2	Ar2299	cobalt transport A
154	6	5.1	124	2	B72628	227	6	5.1	206	2	AC1997	hypothetical prote
155	6	5.1	124	2	AS2389	228	6	5.1	207	2	D81999	glucose inhibited
156	6	5.1	127	2	S01397	229	6	5.1	207	2	C81327	glucose inhibited
157	6	5.1	127	2	A54670	230	6	5.1	208	1	E69797	conserved hypothet
158	6	5.1	130	2	C70971	231	6	5.1	208	2	C96773	hypothetical prote
159	6	5.1	134	2	F97475	232	6	5.1	211	1	B70738	hypothetical prote
160	6	5.1	139	2	C70660	233	6	5.1	212	2	AG0854	probable sugar ald
161	6	5.1	139	2	H81886	234	6	5.1	212	2	B69493	hypothetical prote
162	6	5.1	142	2	S45924	235	6	5.1	214	2	S27449	Sc14 protein - bra
163	6	5.1	142	2	AF1517	236	6	5.1	215	2	A83163	probable carboxyle
164	6	5.1	144	2	H84292	237	6	5.1	217	2	B83604	hypothetical prote
165	6	5.1	144	2	D70196	238	6	5.1	218	2	T49885	peptide methionine
166	6	5.1	146	2	S63582	239	6	5.1	218	2	E84335	hypothetical prote
167	6	5.1	146	2	C70314	240	6	5.1	219	2	C83454	probable transcrip
168	6	5.1	150	2	A83374	241	6	5.1	221	2	C95009	potassium uptake p
169	6	5.1	151	2	AF0887	242	6	5.1	224	2	AH2116	hypothetical prote
170	6	5.1	151	2	F84070	243	6	5.1	226	2	AI2668	hypothetical prote
171	6	5.1	151	2	G72647	244	6	5.1	226	2	JC5327	adhesin complex 25
172	6	5.1	152	2	H64243	245	6	5.1	227	2	AH0683	hypothetical prote
173	6	5.1	153	2	F85642	246	6	5.1	228	2	G97880	hypothetical prote
174	6	5.1	153	2	A90782	247	6	5.1	230	2	D83441	two-component resp
175	6	5.1	154	2	S77145	248	6	5.1	230	2	T36672	membrane-spanning

249	6	5.1	232	2	D84006	lytic transglycosy	322	6	5.1	281	2	C54241	hypothetical prote
250	6	5.1	234	2	T53302	hypothetical prote	323	6	5.1	282	2	H75250	Oxra-related prote
251	6	5.1	236	2	T06462	glutathione peroxi	324	6	5.1	283	2	A64999	hypothetical prote
252	6	5.1	237	2	T05973	permatin homolog P	325	6	5.1	285	2	H69369	branched-chain ami
253	6	5.1	237	2	B64643	hypothetical prote	326	6	5.1	289	2	A87534	carboxylesterase f
254	6	5.1	237	2	T70606	probable pknM prot	327	6	5.1	291	2	AB3235	nitrotriacetate
255	6	5.1	237	2	C42606	orfB 3' to orf405	328	6	5.1	293	2	T47163	hypothetical prote
256	6	5.1	238	2	AC2874	chitoooligosacchari	329	6	5.1	295	2	A87509	hypothetical prote
257	6	5.1	238	2	D97650	conserved hypoteth	330	6	5.1	295	2	E97597	hypothetical prote
258	6	5.1	238	2	T31218	hypothetical prote	331	6	5.1	296	2	T05110	hypothetical prote
259	6	5.1	238	2	G73379	conserved hypoteth	332	6	5.1	297	2	G83056	conserved hypoteth
260	6	5.1	239	2	G82383	oxidoreductase, sh	333	6	5.1	297	2	AG2955	hypothetical prote
261	6	5.1	239	2	JC7768	blue fluorescent p	334	6	5.1	297	2	F98327	hypothetical prote
262	6	5.1	239	2	T34950	hypothetical prote	335	6	5.1	300	2	H84310	hypothetical prote
263	6	5.1	239	2	F86738	hypothetical prote	336	6	5.1	300	2	B83100	inhibitor of chrom
264	6	5.1	240	2	A41797	ig light chain - s	337	6	5.1	301	2	S75531	transcription regu
265	6	5.1	240	2	H84197	hypothetical prote	338	6	5.1	301	2	C72616	hypothetical prote
266	6	5.1	240	2	T33165	hypothetical prote	339	6	5.1	302	2	T35246	probable integral
267	6	5.1	241	2	C81971	hypothetical prote	340	6	5.1	302	2	S27846	hypothetical prote
268	6	5.1	241	2	H81026	conserved hypoteth	341	6	5.1	303	1	S75782	methanol dehydroge
269	6	5.1	242	2	G75607	phosphoesterase-re	342	6	5.1	303	2	G70756	hypothetical prote
270	6	5.1	243	2	F69553	enoyl-CoA hydratase	343	6	5.1	303	2	T36509	probable molybdopt
271	6	5.1	245	2	G75422	hypothetical prote	344	6	5.1	303	2	D87352	conserved hypoteth
272	6	5.1	246	2	D64322	ribosomal protein	345	6	5.1	305	2	C82524	pyrroline-5-carbox
273	6	5.1	247	2	AD0757	cobalamin (5'-phos	346	6	5.1	305	2	A40573	ciathrin heavy cha
274	6	5.1	247	2	S64789	hypothetical prote	347	6	5.1	305	2	F72574	probable cytochrom
275	6	5.1	247	2	S22800	terminase gPM - ph	348	6	5.1	306	2	AC2419	ornithine carbamoy
276	6	5.1	248	2	T46903	hypothetical prote	349	6	5.1	306	2	T46757	lipoprotein lmb [v
277	6	5.1	248	2	E70600	hypothetical prote	350	6	5.1	307	1	C70112	conserved hypoteth
278	6	5.1	248	2	T33034	hypothetical prote	351	6	5.1	308	1	S22931	ubiquinol-cytochro
279	6	5.1	249	2	F71311	probable triosepho	352	6	5.1	308	2	S22919	ubiquinol-cytochro
280	6	5.1	249	2	T35724	cobalt transport i	353	6	5.1	308	2	S22921	ubiquinol-cytochro
281	6	5.1	252	2	AE1288	methionine aminope	354	6	5.1	308	2	S22925	ubiquinol-cytochro
282	6	5.1	252	2	AD1660	methionine aminope	355	6	5.1	308	2	S22929	ubiquinol-cytochro
283	6	5.1	253	2	A89955	protein K04F1.1 [i	356	6	5.1	308	2	S22924	ubiquinol-cytochro
284	6	5.1	254	2	B33954	hypothetical prote	357	6	5.1	308	2	S22927	ubiquinol-cytochro
285	6	5.1	255	2	AB1099	1-acyl-sn-glycerol	358	6	5.1	308	2	S22930	ubiquinol-cytochro
286	6	5.1	255	2	AB1842	1-acylglycerol-3-p	359	6	5.1	308	2	S22920	ubiquinol-cytochro
287	6	5.1	257	1	S22363	guFA protein homol	360	6	5.1	308	2	S22932	ubiquinol-cytochro
288	6	5.1	257	2	AF0890	probable membrane	361	6	5.1	308	2	S22928	ubiquinol-cytochro
289	6	5.1	257	2	H91119	guFA protein homol	362	6	5.1	308	2	S22926	ubiquinol-cytochro
290	6	5.1	257	2	G85964	guFA protein homol	363	6	5.1	308	2	S22922	ubiquinol-cytochro
291	6	5.1	257	2	S16865	gene f41 protein -	364	6	5.1	308	2	S22923	ubiquinol-cytochro
292	6	5.1	259	2	C87287	inositol monophosp	365	6	5.1	308	2	D95932	probable sugar upt
293	6	5.1	259	2	T70615	hypothetical prote	366	6	5.1	308	2	AB0343	probable phosphati
294	6	5.1	259	2	S77575	oligopeptide trans	367	6	5.1	310	2	A71374	probable hypotethi
295	6	5.1	260	2	E75592	probable chromosom	368	6	5.1	311	2	A95932	probable sugar kin
296	6	5.1	261	2	D70871	hypothetical prote	369	6	5.1	312	2	C70475	cytochrome c bioge
297	6	5.1	261	2	T43579	type III secretion	370	6	5.1	312	2	T35111	probable tRNA delt
298	6	5.1	262	2	B85651	hypothetical prote	371	6	5.1	312	2	B64072	spermidine/putresc
299	6	5.1	262	2	G90790	hypothetical prote	372	6	5.1	312	2	C87562	conserved hypoteth
300	6	5.1	262	2	H75377	conserved hypoteth	373	6	5.1	312	2	G97382	probable ATP-bindi
301	6	5.1	262	2	AD2819	conserved hypoteth	374	6	5.1	312	2	AF2600	hypothetical prote
302	6	5.1	264	2	H83170	conserved hypoteth	375	6	5.1	313	2	D98251	hypothetical prote
303	6	5.1	265	2	S72778	hypothetical prote	376	6	5.1	313	2	AH3034	lysophospholipase
304	6	5.1	266	2	T31217	transcription regu	377	6	5.1	313	2	D96028	probable transcrip
305	6	5.1	267	2	G97450	ABC transporter, A	378	6	5.1	316	1	NJBYM1	toxin M1-1 precurs
306	6	5.1	267	2	S74415	hypothetical prote	379	6	5.1	316	2	B83658	phosphoribosyl pyr
307	6	5.1	267	2	I40327	baf protein - Bord	380	6	5.1	316	2	H95985	probable transcrip
308	6	5.1	270	2	S62760	cytochrome-c oxida	381	6	5.1	317	2	H82785	dolichol-phosphate
309	6	5.1	270	2	H70690	hypothetical prote	382	6	5.1	318	2	T45397	FixB [imported] -
310	6	5.1	271	2	T11590	hypothetical prote	383	6	5.1	318	2	T36438	probable membrane
311	6	5.1	271	2	T37222	probable secreted	384	6	5.1	319	2	B75495	conserved hypoteth
312	6	5.1	274	2	F70703	probable uspB prot	385	6	5.1	320	2	I51030	ubiquinol-cytochro
313	6	5.1	274	2	T35062	probable citrate 1	386	6	5.1	320	2	S69547	transcription init
314	6	5.1	274	2	E82179	ABC transporter, p	387	6	5.1	321	2	G98067	sucrose regulon re
315	6	5.1	276	2	D87407	conserved hypoteth	388	6	5.1	321	2	A95201	sucrose operon rep
316	6	5.1	276	2	B83721	oligopeptide ABC t	389	6	5.1	321	2	G82752	cobalt-zinc-cadmium
317	6	5.1	277	2	H84314	cytochrome aa3 con	390	6	5.1	322	2	T20272	hypothetical prote
318	6	5.1	280	2	H95113	conserved hypoteth	391	6	5.1	323	2	A72508	probable cobalamin
319	6	5.1	280	2	A97983	conserved hypoteth	392	6	5.1	324	2	F75631	iron ABC transport
320	6	5.1	281	2	C95932	probable sugar upt	393	6	5.1	325	2	T31474	hypothetical prote
321	6	5.1	281	2	T05522	hypothetical prote	394	6	5.1	325	2	G96718	unknown protein, 5

395	6	5.1	326	2	AG1688	protein-tyrosine/s	468	6	5.1	371	2	B98308	3-isopropylmalate
396	6	5.1	327	2	T44255	probable amino aci	469	6	5.1	371	2	T31874	hypothetical prote
397	6	5.1	327	2	D11491	hypothetical prote	470	6	5.1	371	2	T51865	vasopressin V2 rec
398	6	5.1	327	2	D30275	hypothetical prote	471	6	5.1	373	2	AD0262	probable membrane
399	6	5.1	328	2	D81650	conserved hypotet	472	6	5.1	374	2	S48639	fructose-bisphosph
400	6	5.1	329	2	B82644	sugar-phosphate de	473	6	5.1	374	2	P87596	hypothetical prote
401	6	5.1	329	2	H70744	hypothetical prote	474	6	5.1	375	2	AD0677	probable aldehyde-
402	6	5.1	329	2	T38355	zinc finger protei	475	6	5.1	375	2	T46378	hypothetical prote
403	6	5.1	331	2	A11365	probable lamda CI	476	6	5.1	377	2	T11361	ubiquinol-cytochro
404	6	5.1	331	2	A13267	peptidyl-prolyl ci	477	6	5.1	377	2	T36246	ubiquinol-cytochro
405	6	5.1	332	2	A11355	low-affinity inorg	478	6	5.1	378	1	S17412	probable glycolate
406	6	5.1	332	2	A11726	low-affinity inorg	479	6	5.1	378	2	S39007	ubiquinol-cytochro
407	6	5.1	333	2	B64085	glpX protein - Hae	480	6	5.1	378	2	B75547	iron-sulfur cofact
408	6	5.1	334	2	S54438	hemin permease [va	481	6	5.1	379	1	CBBO	ubiquinol-cytochro
409	6	5.1	335	2	T50601	endo-1,4-beta-xyla	482	6	5.1	379	1	S17405	ubiquinol-cytochro
410	6	5.1	335	2	A40038	MHC class I histoc	483	6	5.1	379	1	S17406	ubiquinol-cytochro
411	6	5.1	335	2	D70874	probable membrane	484	6	5.1	379	1	S17407	ubiquinol-cytochro
412	6	5.1	336	2	H72618	hypothetical prote	485	6	5.1	379	1	S17408	ubiquinol-cytochro
413	6	5.1	336	2	T32902	hypothetical prote	486	6	5.1	379	1	S17409	ubiquinol-cytochro
414	6	5.1	337	1	Q8ECH3	probable dehydroge	487	6	5.1	379	1	S17410	ubiquinol-cytochro
415	6	5.1	337	2	C31029	probable PTS syste	488	6	5.1	379	1	S17411	ubiquinol-cytochro
416	6	5.1	337	2	D85873	probable PTS syste	489	6	5.1	379	1	S17413	ubiquinol-cytochro
417	6	5.1	337	2	A10649	probable regulator	490	6	5.1	379	1	S17414	ubiquinol-cytochro
418	6	5.1	337	2	H85703	Hnr protein [impor	491	6	5.1	379	1	S17417	ubiquinol-cytochro
419	6	5.1	337	2	A36871	37K regulator resp	492	6	5.1	379	1	S17418	ubiquinol-cytochro
420	6	5.1	337	2	A30846	Hnr protein [impor	493	6	5.1	379	1	S17419	ubiquinol-cytochro
421	6	5.1	338	2	A10284	probable response	494	6	5.1	379	1	S17420	ubiquinol-cytochro
422	6	5.1	339	2	S62596	ubiquinol-cytochro	495	6	5.1	379	1	S26163	ubiquinol-cytochro
423	6	5.1	339	2	T34925	ABC transporter in	496	6	5.1	379	1	S33572	ubiquinol-cytochro
424	6	5.1	340	2	JN0527	tcpg protein precu	497	6	5.1	379	1	S41832	ubiquinol-cytochro
425	6	5.1	340	2	JN0525	toxin co-regulated	498	6	5.1	379	1	S41833	ubiquinol-cytochro
426	6	5.1	342	2	S23438	hypothetical prote	499	6	5.1	379	1	S41834	ubiquinol-cytochro
427	6	5.1	343	2	S62704	ubiquinol-cytochro	500	6	5.1	379	1	S41847	ubiquinol-cytochro
428	6	5.1	343	2	F83126	ferric enterobacti	501	6	5.1	379	1	S43261	ubiquinol-cytochro
429	6	5.1	343	2	T26784	hypothetical prote	502	6	5.1	379	1	S43262	ubiquinol-cytochro
430	6	5.1	345	2	C33830	cation efflux syst	503	6	5.1	379	1	S43263	ubiquinol-cytochro
431	6	5.1	345	2	A82348	probable ADP-hepto	504	6	5.1	379	1	S43264	ubiquinol-cytochro
432	6	5.1	346	2	A22991	ABC transporter, m	505	6	5.1	379	1	S43265	ubiquinol-cytochro
433	6	5.1	346	2	E98292	ribose ABC transpo	506	6	5.1	379	1	S43266	ubiquinol-cytochro
434	6	5.1	346	2	D83340	hypothetical prote	507	6	5.1	379	1	S43267	ubiquinol-cytochro
435	6	5.1	347	2	A12974	tartrate dehydroge	508	6	5.1	379	1	S43268	ubiquinol-cytochro
436	6	5.1	347	2	T02280	hypothetical prote	509	6	5.1	379	1	S43269	ubiquinol-cytochro
437	6	5.1	348	2	T14206	NADH2 dehydrogenas	510	6	5.1	379	1	S43270	ubiquinol-cytochro
438	6	5.1	348	2	E87731	protein F32B5.1 [i	511	6	5.1	379	2	D90617	cytochrome b [impo
439	6	5.1	348	2	T29876	cobW protein - Rho	512	6	5.1	379	2	S58461	ubiquinol-cytochro
440	6	5.1	348	2	T03530	iron(III) ABC tran	513	6	5.1	379	2	D90621	cytochrome b [impo
441	6	5.1	351	2	A82429	probable trbB prot	514	6	5.1	379	2	E58889	ubiquinol-cytochro
442	6	5.1	352	2	D70788	MoxR-related prote	515	6	5.1	379	2	I48133	ubiquinol-cytochro
443	6	5.1	354	2	D75460	prephenate dehydro	516	6	5.1	379	2	I48132	ubiquinol-cytochro
444	6	5.1	354	2	A86843	hypothetical prote	517	6	5.1	379	2	I48135	ubiquinol-cytochro
445	6	5.1	356	2	E83192	hypothetical prote	518	6	5.1	379	2	I48180	ubiquinol-cytochro
446	6	5.1	359	2	C59369	hypothetical prote	519	6	5.1	379	2	I49399	ubiquinol-cytochro
447	6	5.1	360	2	E72290	branched chain ami	520	6	5.1	379	2	I49400	ubiquinol-cytochro
448	6	5.1	360	2	S74638	alanine dehydrogen	521	6	5.1	379	2	T11453	ubiquinol-cytochro
449	6	5.1	361	2	E75610	conserved hypotet	522	6	5.1	379	2	T11152	ubiquinol-cytochro
450	6	5.1	362	2	D82644	sugar-phosphate de	523	6	5.1	379	2	T11505	ubiquinol-cytochro
451	6	5.1	362	2	T36079	hypothetical prote	524	6	5.1	379	2	S58455	ubiquinol-cytochro
452	6	5.1	363	2	AD2100	alanine dehydrogen	525	6	5.1	379	2	S58454	ubiquinol-cytochro
453	6	5.1	363	2	T34957	probable phospho-N	526	6	5.1	379	2	T11414	ubiquinol-cytochro
454	6	5.1	364	2	G82734	acetylornithine de	527	6	5.1	379	2	S58451	ubiquinol-cytochro
455	6	5.1	364	2	S43574	COSB5.3 protein (c	528	6	5.1	379	2	S58452	ubiquinol-cytochro
456	6	5.1	364	2	C87292	conserved hypotet	529	6	5.1	379	2	S58457	ubiquinol-cytochro
457	6	5.1	365	2	T48135	transcription nega	530	6	5.1	379	2	S58466	ubiquinol-cytochro
458	6	5.1	366	2	F72062	hypothetical prote	531	6	5.1	379	2	S58465	ubiquinol-cytochro
459	6	5.1	366	2	A86561	Ct449 hypothetical	532	6	5.1	379	2	S58449	ubiquinol-cytochro
460	6	5.1	367	2	S19172	cytochrome P450 2B	533	6	5.1	379	2	E58851	ubiquinol-cytochro
461	6	5.1	368	2	AD0599	probable inner mem	534	6	5.1	379	2	T11375	ubiquinol-cytochro
462	6	5.1	369	2	A11079	conserved hypotet	535	6	5.1	379	2	T11869	ubiquinol-cytochro
463	6	5.1	369	2	E75620	hypothetical prote	536	6	5.1	379	2	T11259	ubiquinol-cytochro
464	6	5.1	369	2	G72730	hypothetical prote	537	6	5.1	379	2	T11349	ubiquinol-cytochro
465	6	5.1	370	2	AC1272	alanine dehydrogen	538	6	5.1	379	2	T11492	ubiquinol-cytochro
466	6	5.1	370	2	AE1634	alanine dehydrogen	539	6	5.1	379	2	S58462	ubiquinol-cytochro
467	6	5.1	371	2	T10635	3-hydroxyisobutyla	540	6	5.1	379	2	T10998	ubiquinol-cytochro

541	6	5.1	379	2	D90625	cytochrome b (impo	614	392	2	AG0226	probable exported
542	6	5.1	379	2	D90615	cytochrome b (impo	615	393	2	H75444	branched-chain ami
543	6	5.1	379	2	T11178	ubiquinol-cytochro	616	394	2	B87019	hypothetical prote
544	6	5.1	379	2	T11530	ubiquinol-cytochro	617	396	2	A85619	aspartate aminotra
545	6	5.1	379	2	D90613	cytochrome b (impo	618	396	2	C90755	aspartate aminotra
546	6	5.1	379	2	D90627	cytochrome b (impo	619	396	2	G71309	probable glutamate
547	6	5.1	379	2	D90619	cytochrome b (impo	620	396	2	D72648	hypothetical prote
548	6	5.1	379	2	T11401	ubiquinol-cytochro	621	398	1	S35473	ubiquinol-cytochro
549	6	5.1	379	2	S58057	ubiquinol-cytochro	622	399	2	B87732	protein W10C8.5 [i
550	6	5.1	379	2	S58085	ubiquinol-cytochro	623	399	2	AH2542	hypothetical prote
551	6	5.1	379	2	S58456	ubiquinol-cytochro	624	399	2	C95943	probable choline u
552	6	5.1	379	2	S58459	ubiquinol-cytochro	625	400	2	T03460	probable leucine/i
553	6	5.1	379	2	S58464	ubiquinol-cytochro	626	401	2	C88571	protein C05B5.3 [i
554	6	5.1	379	2	S58458	ubiquinol-cytochro	627	402	2	T34715	probable ornithine
555	6	5.1	379	2	A53077	ubiquinol-cytochro	628	402	2	T15677	probable ornithine
556	6	5.1	379	2	S58447	ubiquinol-cytochro	629	403	2	B43260	phosphoglycerate k
557	6	5.1	379	2	S58448	ubiquinol-cytochro	630	403	2	D82076	trna nucleotidyltr
558	6	5.1	379	2	S58450	ubiquinol-cytochro	631	403	2	F82076	general secretion
559	6	5.1	379	2	S58460	ubiquinol-cytochro	632	405	1	CBQFR	ubiquinol-cytochro
560	6	5.1	379	2	C84577	probable nucleosom	633	405	1	AI0006	DNA/pantothenate m
561	6	5.1	379	2	I48134	ubiquinol-cytochro	634	405	2	G69499	adenosylhomocyste
562	6	5.1	380	1	CBHU	ubiquinol-cytochro	635	406	2	E72545	hypothetical prote
563	6	5.1	380	1	CBXL	ubiquinol-cytochro	636	407	2	A44374	3-carboxy-cis,cis-
564	6	5.1	380	1	S04840	ubiquinol-cytochro	637	407	2	AC0971	conserved hypotet
565	6	5.1	380	1	S10198	ubiquinol-cytochro	638	407	2	C83589	probable manA prot
566	6	5.1	380	1	S36011	ubiquinol-cytochro	639	408	2	A70594	probable multidrug
567	6	5.1	380	2	D59154	ubiquinol-cytochro	640	410	2	E75290	probable multidrug
568	6	5.1	380	2	T11033	ubiquinol-cytochro	641	410	2	H90312	succinyl-diaminopi
569	6	5.1	380	2	I51374	ubiquinol-cytochro	642	411	2	H70908	hypothetical prote
570	6	5.1	380	2	E58893	ubiquinol-cytochro	643	412	2	C87686	hypothetical prote
571	6	5.1	380	2	T11114	ubiquinol-cytochro	644	413	2	AC1045	probable permease
572	6	5.1	380	2	T11518	ubiquinol-cytochro	645	413	2	AG2407	conserved hypotet
573	6	5.1	380	2	T11845	ubiquinol-cytochro	646	414	2	D87448	site-specific DNA-
574	6	5.1	380	2	T11335	ubiquinol-cytochro	647	414	2	P84393	peptidase, M20/M25
575	6	5.1	380	2	T11466	ubiquinol-cytochro	648	415	2	F84393	threonine synthase
576	6	5.1	380	2	T11204	ubiquinol-cytochro	649	416	2	T37023	probable oxidoredu
577	6	5.1	380	2	T11191	ubiquinol-cytochro	650	417	2	JC4698	divalent cation re
578	6	5.1	380	2	T11113	ubiquinol-cytochro	651	418	2	AI2852	poly(A) polymerase
579	6	5.1	380	2	T09869	ubiquinol-cytochro	652	418	2	B81359	UDP-N-acetylglucos
580	6	5.1	380	2	T09959	ubiquinol-cytochro	653	420	2	E75596	probable O-antigen
581	6	5.1	380	2	D90623	cytochrome b (impo	654	421	2	D70868	probable lipQ prot
582	6	5.1	380	2	S42245	ubiquinol-cytochro	655	421	2	T37223	probable secreted
583	6	5.1	380	2	T11086	ubiquinol-cytochro	656	421	2	D95975	hypothetical outer
584	6	5.1	380	2	T11299	ubiquinol-cytochro	657	421	2	AC0346	probable ABC trans
585	6	5.1	380	2	H70590	hypothetical prote	658	421	2	AI3523	glycerol-3-phospha
586	6	5.1	380	2	D84295	hypothetical prote	659	424	2	H69323	translation initia
587	6	5.1	381	1	CBMS	ubiquinol-cytochro	660	425	2	S34449	transcription fact
588	6	5.1	381	2	T11546	ubiquinol-cytochro	661	426	2	JC5086	polytopic cytoplas
589	6	5.1	381	2	T11312	ubiquinol-cytochro	662	427	2	A40735	TGF beta homolog d
590	6	5.1	381	2	T11440	ubiquinol-cytochro	663	427	2	A87280	conserved hypotet
591	6	5.1	381	2	T11776	ubiquinol-cytochro	664	429	2	F71713	glycerol-3-phospha
592	6	5.1	381	2	S68140	ubiquinol-cytochro	665	430	1	A65165	probable voltage g
593	6	5.1	382	1	S33573	ubiquinol-cytochro	666	430	1	A65165	pantothenate metab
594	6	5.1	382	2	S47882	ubiquinol-cytochro	667	430	2	A87008	hypothetical prote
595	6	5.1	382	2	D58930	ubiquinol-cytochro	668	430	2	A90020	preprotein translo
596	6	5.1	382	2	JC2571	cellulase (EC 3.2.	669	430	2	C86040	pantothenate metab
597	6	5.1	383	2	S54213	flagellar biosynth	670	430	2	B91193	pantothenate metab
598	6	5.1	383	2	G75431	probable Na+/H+ an	671	430	2	T34927	probable oxidoredu
599	6	5.1	385	2	S70984	recP protein - Myc	672	432	2	T36482	probable aminopept
600	6	5.1	385	2	S69587	hypothetical prote	673	433	2	F84215	aminopeptidase [im
601	6	5.1	386	2	T11832	ubiquinol-cytochro	674	433	2	A29626	apolipoprotein B -
602	6	5.1	386	2	T11286	ubiquinol-cytochro	675	434	2	AF3481	nicotinate phospho
603	6	5.1	386	2	AD0218	flagellar biosynth	676	435	2	G97629	pola polymerase (p
604	6	5.1	387	1	CBASN	ubiquinol-cytochro	677	435	2	A75305	adenylosuccinate l
605	6	5.1	387	2	T35425	probable fatty aci	678	435	2	E82753	D-amino acid dehyd
606	6	5.1	388	1	S72995	alanine racemase (679	435	2	AC1410	cellobiose phospho
607	6	5.1	388	2	AH1950	carbamoyl phosphat	680	435	2	AB1786	cellobiose phospho
608	6	5.1	389	2	S62597	ubiquinol-cytochro	681	435	2	C95975	probable polyacch
609	6	5.1	390	2	G75341	alanine dehydrogen	682	435	2	A03236	hypothetical prote
610	6	5.1	390	2	AH0260	conserved hypotet	683	436	2	T38812	hypothetical prote
611	6	5.1	391	1	S25183	chloramphenicol re	684	436	2	JC5021	platelet-activatin
612	6	5.1	391	2	F83233	probable MPS trans	685	437	1	A31752	transcription fact
613	6	5.1	391	2	E44490	retrovirus-related	686	437	2	S42111	transcription fact

687	6	5.1	438	2	H83556	probable MFS trans	760	6	5.1	491	2	F96022	conserved hypothet
688	6	5.1	438	2	B64915	probable membrane	761	6	5.1	491	2	AG1384	hypothetical prote
689	6	5.1	438	2	G85784	probable chloride	762	6	5.1	491	2	AI1759	hypothetical prote
690	6	5.1	438	2	B90916	probable chloride	763	6	5.1	493	2	T06031	hexokinase homolog
691	6	5.1	438	2	G87290	major facilitator	764	6	5.1	493	2	S76517	hypothetical prote
692	6	5.1	439	2	T31071	conserved hypothet	765	6	5.1	495	2	E75081	alkaline phosphata
693	6	5.1	440	2	H69989	lipoprotein homolo	766	6	5.1	498	2	T14236	NADH2 dehydrogenas
694	6	5.1	443	1	E64667	multidrug-efflux t	767	6	5.1	498	2	S62626	protein disulfide-
695	6	5.1	443	2	F71848	probable transport	768	6	5.1	503	2	A53956	nicotinic acetylch
696	6	5.1	444	1	E69130	histidine-tRNA lig	769	6	5.1	503	2	S55589	D-nopaline dehydro
697	6	5.1	444	2	D87557	major facilitator	770	6	5.1	503	2	AD3128	choline sulfatase
698	6	5.1	445	2	S67137	hypothetical prote	771	6	5.1	503	2	E98159	probable aldehyde
699	6	5.1	448	2	S45112	transcription fact	772	6	5.1	506	2	T09437	medium-chain-fatty
700	6	5.1	450	2	E98303	hypothetical 49.3K	773	6	5.1	507	2	A99458	probable glucose t
701	6	5.1	450	2	AH2979	hypothetical prote	774	6	5.1	508	2	T05156	major surface glyc
702	6	5.1	451	2	T40758	hypothetical prote	775	6	5.1	508	2	T30547	probable UDP-N-ace
703	6	5.1	452	2	T35729	hypothetical RNA m	776	6	5.1	510	2	A70580	hypothetical prote
704	6	5.1	452	2	AG1293	hypothetical RNA m	777	6	5.1	511	2	B87258	hypothetical prote
705	6	5.1	453	2	AS1665	hypothetical RNA m	778	6	5.1	512	2	D81701	GMP synthase TC044
706	6	5.1	454	2	F75580	probable sugar tra	779	6	5.1	512	2	B86445	unknown protein [i
707	6	5.1	455	2	AF0215	probable 4-hydroxy	780	6	5.1	513	2	T21887	hypothetical prote
708	6	5.1	458	2	C86860	hypothetical prote	781	6	5.1	515	2	AF3524	hypothetical prote
709	6	5.1	460	2	C71884	hypothetical prote	782	6	5.1	517	2	E86270	hypothetical prote
710	6	5.1	461	1	E64630	virulence factor m	783	6	5.1	517	2	H87022	hypothetical prote
711	6	5.1	462	2	D87566	hypothetical prote	784	6	5.1	520	2	AF2370	serine/threonine k
712	6	5.1	463	2	T34841	probable bifunctio	785	6	5.1	523	1	TVFVMT	protein-tyrosine k
713	6	5.1	463	2	H70504	probable GTP-bindin	786	6	5.1	523	1	C70117	probable purH prot
714	6	5.1	464	2	S36501	L2 protein - human	787	6	5.1	524	1	S15619	L2 protein - human
715	6	5.1	465	1	S15626	protein L2 - human	788	6	5.1	524	2	S55097	probable membrane
716	6	5.1	467	2	G82198	RTX toxin transpor	789	6	5.1	526	1	OKFVVR	protein-tyrosine k
717	6	5.1	467	2	B97213	fAD/FMN-containing	790	6	5.1	526	1	TVFV60	protein-tyrosine k
718	6	5.1	468	2	A75145	protein translocas	791	6	5.1	526	1	TVFVR	protein-tyrosine k
719	6	5.1	468	2	D70646	probable acyl-CoA	792	6	5.1	526	2	S15582	protein-tyrosine k
720	6	5.1	468	2	E71184	probable preprotei	793	6	5.1	526	2	S20808	protein-tyrosine k
721	6	5.1	469	2	AC0950	Two-component syst	794	6	5.1	526	2	S26420	protein-tyrosine k
722	6	5.1	469	2	S53024	nitrogen regulatio	795	6	5.1	526	2	B70859	hypothetical prote
723	6	5.1	470	2	D90895	aldehyde dehydroge	796	6	5.1	530	2	H82801	conserved hypothet
724	6	5.1	470	2	D80895	probable aldehyde	797	6	5.1	531	2	T51922	hypothetical prote
725	6	5.1	470	2	D85722	probable aldehyde	798	6	5.1	532	1	A34104	protein-tyrosine k
726	6	5.1	471	2	H87284	phosphate regulon	799	6	5.1	532	1	A34104	protein-tyrosine k
727	6	5.1	472	2	T46217	glutamate-1-semial	800	6	5.1	532	2	A57173	oculocutaneous alb
728	6	5.1	472	2	P82639	resistance protein	801	6	5.1	533	1	TVCHS	protein-tyrosine k
729	6	5.1	474	2	C82737	cysteinyI-tRNA syn	802	6	5.1	534	2	D98224	dipeptide transpor
730	6	5.1	475	2	D64751	amino acid permeas	803	6	5.1	534	2	AC3062	hypothetical prote
731	6	5.1	475	2	S54993	reverse transcript	804	6	5.1	534	2	T34455	hypothetical prote
732	6	5.1	475	2	S54994	reverse transcript	805	6	5.1	535	2	H83324	probable chemotaxi
733	6	5.1	475	2	T35697	arabinofuranosidas	806	6	5.1	536	2	C82433	methyI-accepting c
734	6	5.1	477	2	S07038	OR11 protein - yea	807	6	5.1	536	2	A13078	conserved hypothet
735	6	5.1	478	2	S21455	glutamate-1-semial	808	6	5.1	539	2	A29923	carboxylesterase (
736	6	5.1	478	2	S21454	glutamate-1-semial	809	6	5.1	540	2	G87407	oxidoreductase, GM
737	6	5.1	478	2	G86646	beta-glucosidase (810	6	5.1	540	2	A31584	carboxylesterase (
738	6	5.1	478	2	F90985	mannose-1-phosphat	811	6	5.1	540	2	G75386	probable 2-phospho
739	6	5.1	478	2	A85831	mannose-1-phosphat	812	6	5.1	541	1	A43610	protein-tyrosine k
740	6	5.1	478	2	H64970	mannose-1-phosphat	813	6	5.1	541	2	A82276	aldehyde dehydroge
741	6	5.1	479	2	T93301	hypothetical prote	814	6	5.1	541	2	H83445	probable chemotaxi
742	6	5.1	479	2	B75102	NADH dehydrogenase	815	6	5.1	542	1	TVHUSC	protein-tyrosine k
743	6	5.1	479	2	T50726	hypothetical prote	816	6	5.1	543	1	F64871	oligopeptide-bindin
744	6	5.1	479	2	F84179	hypothetical prote	817	6	5.1	543	2	F85704	hypothetical prote
745	6	5.1	479	2	S18447	variant surface gl	818	6	5.1	543	2	G90846	hypothetical prote
746	6	5.1	480	2	AF0768	mannose-1-phosphat	819	6	5.1	545	2	S52313	protein-tyrosine k
747	6	5.1	480	2	S65290	mannose-1-phosphat	820	6	5.1	546	2	S52314	hypothetical prote
748	6	5.1	480	2	H82506	hypothetical prote	821	6	5.1	547	2	T30269	hypothetical prote
749	6	5.1	481	2	T07034	glutamate-1-semial	822	6	5.1	548	2	B84306	hypothetical prote
750	6	5.1	482	2	H69392	4-hydroxybutyrate	823	6	5.1	548	2	A75357	hypothetical prote
751	6	5.1	485	2	AD0041	rhamnulokinase (EC	824	6	5.1	548	2	H70788	probable peptidatr
752	6	5.1	485	2	T13479	hypothetical prote	825	6	5.1	549	2	JX0054	carboxylesterase (
753	6	5.1	487	2	S61993	probable membrane	826	6	5.1	551	2	H98207	hypothetical prote
754	6	5.1	489	2	F83109	probable phosphate	827	6	5.1	551	2	AI1829	hypothetical prote
755	6	5.1	491	1	O4RTPB	cytochrome P450 2B	828	6	5.1	552	2	T06491	beta-fructofuranos
756	6	5.1	491	1	O4RTP2	cytochrome P450 2B	829	6	5.1	555	2	F72555	probable molybdenu
757	6	5.1	491	2	B86096	xylose-proton symp	830	6	5.1	555	2	AG3432	chloride channel p
758	6	5.1	491	2	F91255	xylose-proton symp	831	6	5.1	557	1	TVFV82	protein-tyrosine k
759	6	5.1	491	2	A26430	xylose transport p	832	6	5.1	557	2	E86106	yidB protein [slmi

979	6	5.1	998	2	GB3022	1052	6	5.1	1765	2	T42388	sodium channel alp
980	6	5.1	1000	2	AB3467	1053	6	5.1	1772	2	A4532	major merizoite su
981	6	5.1	1015	2	JC5062	1054	6	5.1	1784	2	B86921	polyketide synthas
982	6	5.1	1015	2	JC5263	1055	6	5.1	1795	2	F97713	190K antigen precu
983	6	5.1	1017	2	T30542	1056	6	5.1	1867	2	T38348	probable i,3-beta-
984	6	5.1	1019	2	A83613	1057	6	5.1	1900	2	AG2391	serine/threonine k
985	6	5.1	1021	2	H75423	1058	6	5.1	1911	2	T43048	calcium channel al
986	6	5.1	1022	2	T30543	1059	6	5.1	2076	2	SI5999	fatty-acyl-CoA syn
987	6	5.1	1029	2	F96602	1060	6	5.1	2124	2	H83357	probable non-ribos
988	6	5.1	1029	2	C97665	1061	6	5.1	2140	2	T18543	probable cell-adhe
989	6	5.1	1029	2	AF2889	1062	6	5.1	2354	2	T13288	mei-1 protein - f
990	6	5.1	1034	2	T30331	1063	6	5.1	2957	2	T33152	hypothetical prote
991	6	5.1	1041	2	T31097	1064	6	5.1	3005	2	S33642	homeotic protein z
992	6	5.1	1051	2	B27672	1065	6	5.1	3033	1	GNWJ08	genome polyprotein
993	6	5.1	1053	2	S58883	1066	6	5.1	3104	2	S20473	fatty-acid synthas
994	6	5.1	1057	2	S09112	1067	6	5.1	3133	2	S52093	hemocytin - silkwo
995	6	5.1	1064	2	H97657	1068	6	5.1	3161	2	T30342	protein HMWPI - Ye
996	6	5.1	1064	2	A42881	1069	6	5.1	3163	2	AB0233	versiniaabactin bio
997	6	5.1	1073	2	H82300	1070	6	5.1	3163	2	T17440	probable polyketid
998	6	5.1	1086	2	G02257	1071	6	5.1	4006	2	T09070	probable tenascin
999	6	5.1	1090	2	JC1421	1072	6	5.1	4096	2	A57099	DNA-activated prot
1000	6	5.1	1091	2	AF2953	1073	6	5.1	4128	2	JC6306	protein kinase (EC
1001	6	5.1	1091	2	G98329	1074	6	5.1	7463	2	T36248	CDA peptide synthe
1002	6	5.1	1105	2	B64973	1075	5	4.2	9	2	PW0002	chlorophyll a/b-bi
1003	6	5.1	1108	2	A48508	1076	5	4.2	13	2	A38929	glutathione peroxi
1004	6	5.1	1165	2	A70423	1077	5	4.2	14	1	QWAVV	mastoparan - yello
1005	6	5.1	1180	2	A11939	1078	5	4.2	14	2	S66234	sperm motility inh
1006	6	5.1	1199	2	T13946	1079	5	4.2	15	2	PA0038	protein QAL00030 -
1007	6	5.1	1204	2	B81947	1080	5	4.2	15	2	A26997	unspecific monooxy
1008	6	5.1	1204	2	R81158	1081	5	4.2	17	2	PA0004	plastocyanin - Ara
1009	6	5.1	1214	2	A62897	1082	5	4.2	20	2	A60728	cytochrome P450 3A
1010	6	5.1	1218	2	A00837	1083	5	4.2	21	2	S78575	protein kinase C i
1011	6	5.1	1218	2	T30293	1084	5	4.2	24	2	B44008	lethal peptide I -
1012	6	5.1	1226	2	S15053	1085	5	4.2	24	2	A44008	lethal peptide II
1013	6	5.1	1227	2	C97033	1086	5	4.2	24	2	D85955	hypothetical prote
1014	6	5.1	1230	2	T22458	1087	5	4.2	25	2	B32351	39K class A flagel
1015	6	5.1	1275	2	I38588	1088	5	4.2	26	2	S59906	gamma-glutamyl tra
1016	6	5.1	1275	2	S65824	1089	5	4.2	28	2	PA0033	cytochrome P450 te
1017	6	5.1	1275	2	B28096	1090	5	4.2	31	2	S53015	DNA-directed RNA p
1018	6	5.1	1278	2	A71609	1091	5	4.2	34	1	LNBOC1	pulmonary surfacta
1019	6	5.1	1280	2	B34087	1092	5	4.2	35	1	LNPGC1	pulmonary surfacta
1020	6	5.1	1283	2	A47377	1093	5	4.2	35	1	LNPGC1	hypothetical prote
1021	6	5.1	1284	2	T13168	1094	5	4.2	35	2	S20042	hypothetical prote
1022	6	5.1	1286	2	A88396	1095	5	4.2	38	2	B82413	hypothetical prote
1023	6	5.1	1333	2	G84542	1096	5	4.2	39	2	I54374	Gene NF2 protein -
1024	6	5.1	1341	2	T17285	1097	5	4.2	40	1	FDFI8G	antifreeze protein
1025	6	5.1	1353	1	AJFPFM	1098	5	4.2	40	2	S04639	dermatan sulfate p
1026	6	5.1	1366	2	T35985	1099	5	4.2	41	2	B24802	cuticle protein 32
1027	6	5.1	1387	2	A97673	1100	5	4.2	41	2	C82793	hypothetical prote
1028	6	5.1	1398	2	C87448	1101	5	4.2	43	2	S49760	extensin - tomato
1029	6	5.1	1405	2	H87230	1102	5	4.2	45	2	S12208	pyruvate dehydroge
1030	6	5.1	1407	2	T28702	1103	5	4.2	45	2	A05163	antifreeze protein
1031	6	5.1	1417	2	T06661	1104	5	4.2	45	2	F64525	hypothetical prote
1032	6	5.1	1440	2	T44872	1105	5	4.2	45	2	B64014	hypothetical prote
1033	6	5.1	1468	2	S88250	1106	5	4.2	46	2	I54375	Gene NF2 protein -
1034	6	5.1	1475	2	T68399	1107	5	4.2	47	2	I64800	sepiapterin reduct
1035	6	5.1	1488	2	T02856	1108	5	4.2	50	1	VCBP22	coat protein B - p
1036	6	5.1	1529	2	T16779	1109	5	4.2	51	2	AH0379	conserved hypothet
1037	6	5.1	1537	2	F86509	1110	5	4.2	52	2	T08487	hypothetical prote
1038	6	5.1	1537	2	C81558	1111	5	4.2	52	2	S75006	hypothetical prote
1039	6	5.1	1537	2	H72112	1112	5	4.2	52	2	T15646	hypothetical prote
1040	6	5.1	1584	1	JN0114	1113	5	4.2	52	2	A10472	hypothetical prote
1041	6	5.1	1584	2	T15822	1114	5	4.2	53	2	D72610	hypothetical prote
1042	6	5.1	1602	2	H70984	1115	5	4.2	54	2	S10109	hypothetical prote
1043	6	5.1	1630	2	A53577	1116	5	4.2	54	2	P82545	hypothetical prote
1044	6	5.1	1644	2	T13803	1117	5	4.2	54	2	D97847	hypothetical prote
1045	6	5.1	1647	2	T49412	1118	5	4.2	54	2	AD2043	hypothetical prote
1046	6	5.1	1651	2	JC1340	1119	5	4.2	55	2	S01504	H+-transporting tw
1047	6	5.1	1675	1	LRRTH	1120	5	4.2	55	2	A43896	76K cell surface l
1048	6	5.1	1678	2	S52588	1121	5	4.2	56	2	T30739	hypothetical prote
1049	6	5.1	1681	2	S42369	1122	5	4.2	56	2	A82565	hypothetical prote
1050	6	5.1	1733	2	D70887	1123	5	4.2	57	2	H86797	prophage pi3 prote
1051	6	5.1	1756	2	T07566	1124	5	4.2	58	2	B43934	probable DNA-bind

1125	5	4.2	58	2	G90329	hypothetical prote	1198	5	4.2	82	2	AH0032	probable type III
1126	5	4.2	59	2	C95155	hypothetical prote	1199	5	4.2	83	2	B84268	hypothetical prote
1127	5	4.2	59	2	D84234	hypothetical prote	1200	5	4.2	83	2	B84494	hypothetical prote
1128	5	4.2	59	2	G98021	hypothetical prote	1201	5	4.2	83	2	A36505	oxaloacetate decar
1129	5	4.2	60	2	I69770	MHC RT44 protein -	1202	5	4.2	83	2	T34904	hypothetical prote
1130	5	4.2	61	2	E86084	hypothetical prote	1203	5	4.2	84	1	W4WL	E4 protein - human
1131	5	4.2	62	2	G83425	hypothetical prote	1204	5	4.2	84	2	D70672	hypothetical prote
1132	5	4.2	63	2	T04493	hypothetical prote	1205	5	4.2	84	2	A64344	hypothetical prote
1133	5	4.2	63	2	T22080	hypothetical prote	1206	5	4.2	84	2	A64344	hypothetical prote
1134	5	4.2	64	2	H95275	hypothetical prote	1207	5	4.2	85	2	JC2290	hypothetical prote
1135	5	4.2	65	2	S75236	hypothetical prote	1208	5	4.2	85	2	S14026	hypothetical prote
1136	5	4.2	65	2	A75543	transcription regu	1209	5	4.2	86	2	F87057	hypothetical prote
1137	5	4.2	65	2	AB3394	hypothetical cytos	1210	5	4.2	86	2	F87057	hypothetical prote
1138	5	4.2	66	2	S04145	infected cell prot	1211	5	4.2	87	2	G46449	hypothetical prote
1139	5	4.2	66	2	C69833	hypothetical prote	1212	5	4.2	87	2	S01982	hypothetical prote
1140	5	4.2	67	2	S76580	hypothetical prote	1213	5	4.2	88	2	I59656	MHC class II antig
1141	5	4.2	67	2	E86573	hypothetical prote	1214	5	4.2	88	2	S60188	l1beta-hydroxyster
1142	5	4.2	67	2	H72049	hypothetical prote	1215	5	4.2	88	2	B85731	hypothetical prote
1143	5	4.2	67	2	A71289	probable DNA-dirc	1216	5	4.2	88	2	G81178	conserved hypothet
1144	5	4.2	67	2	H95418	hypothetical prote	1217	5	4.2	88	2	B81926	hypothetical prote
1145	5	4.2	67	2	AB0396	hypothetical prote	1218	5	4.2	88	2	A81897	hypothetical prote
1146	5	4.2	68	2	S18622	hypothetical prote	1219	5	4.2	89	2	I53416	interleukin-8 homo
1147	5	4.2	68	2	AG3335	regulatory protein	1220	5	4.2	89	2	A53497	pre-B-cell growth-
1148	5	4.2	69	2	T44956	hypothetical prote	1221	5	4.2	89	2	T08479	plasmid maintenanc
1149	5	4.2	69	2	C46449	70K heat shock cha	1222	5	4.2	89	2	D90802	probable transcript
1150	5	4.2	69	2	S15956	hypothetical prote	1223	5	4.2	89	2	T45015	hypothetical prote
1151	5	4.2	69	2	T43093	entry exclusion pr	1224	5	4.2	89	2	A70703	hypothetical prote
1152	5	4.2	69	2	AC3576	hypothetical prote	1225	5	4.2	89	2	A43854	major merozoite su
1153	5	4.2	70	2	AG1015	hypothetical membr	1226	5	4.2	90	2	G75317	hypothetical prote
1154	5	4.2	70	2	S34217	conserved hypothet	1227	5	4.2	90	2	A82192	transglycosylase-a
1155	5	4.2	71	2	A36919	acclimation protei	1228	5	4.2	91	2	A98116	hypothetical prote
1156	5	4.2	71	2	PQ0114	cryptophan 2-monoo	1229	5	4.2	91	2	S52275	hypothetical prote
1157	5	4.2	71	2	C53306	hypothetical prote	1230	5	4.2	91	2	A85961	hypothetical prote
1158	5	4.2	71	2	T03417	k1eb protein - pla	1231	5	4.2	92	2	F69390	LSU ribosomal prot
1159	5	4.2	72	1	LHW43	trad protein - Agr	1232	5	4.2	92	2	G84391	hypothetical prote
1160	5	4.2	72	2	AG2722	H+-transporting tw	1233	5	4.2	92	2	T36228	hypothetical prote
1161	5	4.2	73	1	NTKN6G	hypothetical prote	1234	5	4.2	92	2	C89554	conserved hypothet
1162	5	4.2	73	1	VCBPF1	omega-conotoxin GV	1235	5	4.2	93	2	I50809	peptide YY - river
1163	5	4.2	73	1	VCBPF1	coat protein B pre	1236	5	4.2	93	2	I81182	cytokine - mouse
1164	5	4.2	73	1	VCBPM3	coat protein B pre	1237	5	4.2	93	2	F84175	hypothetical prote
1165	5	4.2	73	1	VCBPM3	coat protein B pre	1238	5	4.2	93	2	D69072	hypothetical prote
1166	5	4.2	73	2	H71149	hypothetical prote	1239	5	4.2	93	2	G84065	hypothetical prote
1167	5	4.2	74	1	WSWL1B	hypothetical prote	1240	5	4.2	93	2	F84116	hypothetical prote
1168	5	4.2	74	2	B59135	probable omega-con	1241	5	4.2	94	2	B81071	hypothetical prote
1169	5	4.2	74	2	T22077	hypothetical prote	1242	5	4.2	94	2	AG1891	hypothetical prote
1170	5	4.2	74	2	C84034	transcription regu	1243	5	4.2	95	2	A75133	hypothetical prote
1171	5	4.2	74	2	C61055	E5b protein - huma	1244	5	4.2	95	2	B75507	hypothetical prote
1172	5	4.2	74	2	D81796	exodeoxyribonuclea	1245	5	4.2	96	2	T17828	hypothetical prote
1173	5	4.2	74	2	G81217	exodeoxyribonuclea	1246	5	4.2	96	2	B84154	hypothetical prote
1174	5	4.2	75	2	S50353	hypothetical prote	1247	5	4.2	97	2	S53315	stilbene synthase
1175	5	4.2	75	2	A80698	probable pathogeni	1248	5	4.2	97	2	T29101	hypothetical prote
1176	5	4.2	75	2	H90886	hypothetical prote	1249	5	4.2	97	2	S31094	hypothetical prote
1177	5	4.2	75	2	E71348	hypothetical prote	1250	5	4.2	97	2	E86682	prophage p11 prote
1178	5	4.2	75	2	G95891	hypothetical prote	1251	5	4.2	97	2	A86754	prophage p12 prote
1179	5	4.2	76	2	B28572	hypothetical prote	1252	5	4.2	97	2	E83584	hypothetical prote
1180	5	4.2	76	2	T30388	ig heavy chain V r	1253	5	4.2	97	2	D83054	hypothetical prote
1181	5	4.2	77	2	T18001	fibrous body prote	1254	5	4.2	97	2	C75415	hypothetical prote
1182	5	4.2	77	2	H64898	hypothetical prote	1255	5	4.2	97	2	T34765	small hydrophobic
1183	5	4.2	78	2	G69033	conserved hypothet	1256	5	4.2	98	2	S24920	NADH2 dehydrogenas
1184	5	4.2	78	2	B91094	type III secretion	1257	5	4.2	98	2	S41828	NADH2 dehydrogenas
1185	5	4.2	78	2	F87350	hypothetical prote	1258	5	4.2	98	2	C86517	hypothetical prote
1186	5	4.2	78	2	D82774	hypothetical prote	1259	5	4.2	98	2	F96643	hypothetical prote
1187	5	4.2	78	2	F85939	type III secretion	1260	5	4.2	98	2	T28629	Y4Cb protein - Rhi
1188	5	4.2	79	2	C87490	hypothetical prote	1261	5	4.2	98	2	T39687	hypothetical prote
1189	5	4.2	79	2	T36418	hypothetical prote	1262	5	4.2	98	2	B44366	zinc finger protei
1190	5	4.2	79	2	B30924	hypothetical prote	1263	5	4.2	98	2	D72106	hypothetical prote
1191	5	4.2	79	2	AC0013	hypothetical 9.8K	1264	5	4.2	99	1	CUFO	hypothetical prote
1192	5	4.2	79	2	S61446	cellulase (EC 3.2.	1265	5	4.2	99	1	CUFO	plastocyanin - she
1193	5	4.2	80	2	H75569	hypothetical prote	1266	5	4.2	99	1	CUVM	plastocyanin - file
1194	5	4.2	81	2	C69876	conserved hypothet	1267	5	4.2	99	2	T43238	ribosomal protein
1195	5	4.2	81	2	AC3243	conjugal transfer	1268	5	4.2	99	2	T45405	hypothetical prote
1196	5	4.2	82	2	I61829	gene MHC DO-beta 1	1269	5	4.2	99	2	T32227	hypothetical prote
1197	5	4.2	82	2	G84088	hypothetical prote	1270	5	4.2	99	2	A82610	conserved hypothet

1271	5	4.2	100	2	JH0269	cystatin - avocado	1344	5	4.2	110	2	S50345	hypothetical prote
1272	5	4.2	100	2	E30247	DNA binding protei	1345	5	4.2	110	2	D83985	general stress pro
1273	5	4.2	100	2	A85525	hypothetical prote	1346	5	4.2	110	2	A72652	hypothetical prote
1274	5	4.2	100	2	F90674	hypothetical prote	1347	5	4.2	111	2	S22898	T-cell receptor al
1275	5	4.2	100	2	AH2220	hypothetical prote	1348	5	4.2	111	2	S19415	hypothetical prote
1276	5	4.2	100	2	A75528	conserved hypothet	1349	5	4.2	111	2	B83101	hypothetical prote
1277	5	4.2	100	2	S30171	mercuric ion trans	1350	5	4.2	111	2	D85608	hypothetical prote
1278	5	4.2	101	2	B91023	hypothetical prote	1351	5	4.2	111	2	AH0118	insertion sequence
1279	5	4.2	101	2	C95867	hypothetical prote	1352	5	4.2	111	2	S03714	zein precursor - m
1280	5	4.2	101	2	H64997	hypothetical prote	1353	5	4.2	112	1	HSUR42	histone H2A, sperm
1281	5	4.2	101	2	S64222	probable membrane	1354	5	4.2	112	2	S14355	glutathione transf
1282	5	4.2	101	2	S45387	hypothetical prote	1355	5	4.2	112	2	A37203	lens fiber membran
1283	5	4.2	101	2	T37064	hypothetical prote	1356	5	4.2	112	2	T36117	conserved hypothet
1284	5	4.2	101	2	A71117	hypothetical prote	1357	5	4.2	112	2	A69288	hypothetical prote
1285	5	4.2	101	2	B97800	hypothetical prote	1358	5	4.2	112	2	AF0225	conserved hypothet
1286	5	4.2	102	1	W4WL31	B4 protein - human	1359	5	4.2	113	2	F71670	ribosomal protein
1287	5	4.2	102	2	D37386	hypothetical prote	1360	5	4.2	113	2	S39031	molt-inhibiting ho
1288	5	4.2	102	2	S44164	hypothetical prote	1361	5	4.2	113	2	A85796	hypothetical prote
1289	5	4.2	102	2	S02221	hypothetical prote	1362	5	4.2	113	2	E90947	hypothetical prote
1290	5	4.2	102	2	T36017	probable SecG-like	1363	5	4.2	113	2	G64945	hypothetical prote
1291	5	4.2	103	1	GRKTM1	matrix Gla protein	1364	5	4.2	113	2	A81093	hypothetical prote
1292	5	4.2	103	1	GBBOM	matrix Gla protein	1365	5	4.2	113	2	E81848	hypothetical prote
1293	5	4.2	103	1	GBHUM	matrix Gla protein	1366	5	4.2	113	2	D72718	hypothetical prote
1294	5	4.2	103	1	GERBM1	matrix Gla protein	1367	5	4.2	114	2	T50194	very hypothetical
1295	5	4.2	103	2	AF0795	conserved hypothet	1368	5	4.2	113	2	A83429	hypothetical prote
1296	5	4.2	103	2	B87261	hypothetical prote	1369	5	4.2	113	2	AH0397	probable lipoprote
1297	5	4.2	103	2	S75322	hypothetical prote	1370	5	4.2	113	2	T49533	hypothetical prote
1298	5	4.2	103	2	H90234	hypothetical prote	1371	5	4.2	113	2	T45276	hypothetical prote
1299	5	4.2	103	2	JQ0306	hypothetical 11.4K	1372	5	4.2	113	2	S43435	nuclear receptor p
1300	5	4.2	104	1	GBMSM1	matrix Gla protein	1373	5	4.2	114	2	T11159	NADH2 dehydrogen
1301	5	4.2	104	2	S16310	bclA protein - Ery	1374	5	4.2	114	2	A54663	seminal plasma pro
1302	5	4.2	104	2	AC0162	probable membrane	1375	5	4.2	114	2	H86896	hypothetical prote
1303	5	4.2	104	2	A12730	transcription regu	1376	5	4.2	114	2	G40511	basic transactivat
1304	5	4.2	105	2	T45377	ribosomal protein	1377	5	4.2	115	2	S15064	somatostatin II pr
1305	5	4.2	105	2	F70643	probable ribosomal	1378	5	4.2	115	2	S37264	MHC class II histo
1306	5	4.2	105	2	S11410	hypothetical prote	1379	5	4.2	115	2	S37265	MHC class II histo
1307	5	4.2	105	2	T12848	hypothetical prote	1380	5	4.2	115	2	T45838	MHC class II DR-be
1308	5	4.2	105	2	G95371	hypothetical prote	1381	5	4.2	115	2	A23329	Ly-5-8 glycoprotei
1309	5	4.2	105	2	A56897	glutamate-5-semial	1382	5	4.2	115	2	B25120	major merozoite su
1310	5	4.2	105	2	GB7591	hypothetical prote	1383	5	4.2	115	2	A86853	hypothetical prote
1311	5	4.2	106	2	I40030	probable dioxygena	1384	5	4.2	115	2	D64710	hypothetical prote
1312	5	4.2	106	2	S25604	hypothetical prote	1385	5	4.2	115	2	A71809	hypothetical prote
1313	5	4.2	106	2	C93584	conserved hypothet	1386	5	4.2	115	2	T51382	hypothetical prote
1314	5	4.2	106	2	A33447	hypothetical prote	1387	5	4.2	115	2	B71029	hypothetical prote
1315	5	4.2	106	2	C36925	cbbR 3'-region hyp	1388	5	4.2	116	1	G1MS10	ig heavy chain pre
1316	5	4.2	106	2	C87523	filio protein [impo	1389	5	4.2	116	1	R5HSH9	ribosomal protein
1317	5	4.2	106	2	D87367	parvalbumin beta [1390	5	4.2	116	2	T10739	thiorodoxin - comm
1318	5	4.2	107	1	PVPK2	cytochrome c [vali	1391	5	4.2	116	2	A33932	ig mu chain precu
1319	5	4.2	107	1	CCHB	probable auxin-ind	1392	5	4.2	116	2	B83114	50S ribosomal prot
1320	5	4.2	107	2	T05296	hypothetical prote	1393	5	4.2	116	2	T02803	probable membrane
1321	5	4.2	107	2	S64345	probable small sec	1394	5	4.2	116	2	D84285	S-D-lactoylglutath
1322	5	4.2	107	2	T35523	conserved hypothet	1395	5	4.2	116	2	C69201	ribosomal protein
1323	5	4.2	107	2	A87637	hypothetical prote	1396	5	4.2	117	1	R5EC18	ribosomal protein
1324	5	4.2	107	2	G81080	hypothetical prote	1397	5	4.2	117	1	T43941	NADH2 dehydrogen
1325	5	4.2	108	1	PVHK	parvalbumin beta -	1398	5	4.2	117	2	T11123	NADH2 dehydrogen
1326	5	4.2	108	1	C69151	conserved hypothet	1399	5	4.2	117	2	S62703	ig kappa chain v r
1327	5	4.2	108	2	F86741	conserved hypothet	1400	5	4.2	117	2	S43528	ribosomal protein
1328	5	4.2	108	2	C64463	hypothetical prote	1401	5	4.2	117	2	JC2282	ribosomal protein
1329	5	4.2	108	2	T09703	pop3 protein - wes	1402	5	4.2	117	2	A11007	50S ribosomal chai
1330	5	4.2	108	2	A75625	hypothetical prote	1403	5	4.2	117	2	A91150	50S ribosomal subu
1331	5	4.2	108	2	A95904	probable response	1404	5	4.2	117	2	B85995	50S ribosomal subu
1332	5	4.2	108	2	T22870	hypothetical prote	1405	5	4.2	117	2	F82057	ribosomal protein
1333	5	4.2	109	1	PVTFB3	parvalbumin beta I	1406	5	4.2	117	2	JC5753	ribosomal protein
1334	5	4.2	109	2	S60468	NADH2 dehydrogen	1407	5	4.2	117	2	C64094	ribosomal protein
1335	5	4.2	109	2	E69849	hypothetical prote	1408	5	4.2	117	2	AB0028	50S ribosomal prot
1336	5	4.2	109	2	C83197	membrane protein G	1409	5	4.2	117	2	B84269	50S ribosomal prot
1337	5	4.2	109	2	B72542	hypothetical prote	1410	5	4.2	117	2	H83969	small basic protei
1338	5	4.2	109	2	A87408	transcription regu	1411	5	4.2	117	2	AF0919	probable membrane
1339	5	4.2	109	2	S29655	juvenile-hormone e	1412	5	4.2	117	2	H84408	hypothetical prote
1340	5	4.2	110	2	S57660	argininosuccinate	1413	5	4.2	117	2	AB3298	hypothetical prote
1341	5	4.2	110	2	B72496	hypothetical prote	1414	5	4.2	118	1	DNOB03	NADH2 dehydrogen
1342	5	4.2	110	2	A71454	hypothetical prote	1415	5	4.2	118	1	DNWTU3	NADH2 dehydrogen
1343	5	4.2	110	2	S38067	hypothetical prote	1416	5	4.2	118	1	DNZMU3	NADH2 dehydrogen

1417	5	4.2	118	2	S71077	NADH2 dehydrogenas	1490	5	4.2	124	2	A70692	hypothetical prote
1418	5	4.2	118	2	S71081	NADH2 dehydrogenas	1491	5	4.2	124	2	JC4849	H+-exporting ATPase
1419	5	4.2	118	2	S25944	NADH2 dehydrogenas	1492	5	4.2	124	2	S55042	11 kDa proteinase
1420	5	4.2	118	2	S71079	NADH2 dehydrogenas	1493	5	4.2	124	2	T37059	hypothetical prote
1421	5	4.2	118	2	T09501	probable NADH2 deh	1494	5	4.2	124	2	AH3226	conserved hypotet
1422	5	4.2	118	2	T09519	NADH2 dehydrogenas	1495	5	4.2	124	2	I69801	rhaR protein - Esc
1423	5	4.2	118	2	S70000	NADH2 dehydrogenas	1496	5	4.2	124	2	A30043	hypothetical prote
1424	5	4.2	118	2	S37201	anti-DNA autoantib	1497	5	4.2	124	2	F82655	trophoblast-epeti
1425	5	4.2	118	2	PL0200	Ig heavy chain V r	1498	5	4.2	124	2	AH1801	efflux protein (tr
1426	5	4.2	118	2	S47024	ribosomal protein	1499	5	4.2	124	2	AE3047	hypothetical prote
1427	5	4.2	118	2	E97289	uncharacterized co	1500	5	4.2	125	1	HSUR9M	histone H2A, gonad
1428	5	4.2	118	2	T31183	hypothetical prote	1501	5	4.2	125	1	HSUR9P	histone H2A, gonad
1429	5	4.2	118	2	T71235	hypothetical prote	1502	5	4.2	125	1	TS0866	probable thioredox
1430	5	4.2	118	2	T06550	pollen allergen ho	1503	5	4.2	125	2	I50498	urotensin II-alpha
1431	5	4.2	118	2	H87630	transcription regu	1504	5	4.2	125	2	T07992	ribosomal protein
1432	5	4.2	118	2	A75566	HesB/YadR/YfjF fam	1505	5	4.2	126	2	T24181	hypothetical prote
1433	5	4.2	118	2	G82700	conserved hypotet	1506	5	4.2	126	2	S65807	transcription acti
1434	5	4.2	118	2	C95398	hypothetical prote	1507	5	4.2	126	2	S75768	hypothetical prote
1435	5	4.2	119	1	CUPSAM	amcyanin precursor	1508	5	4.2	126	2	AB2905	conserved hypotet
1436	5	4.2	119	2	E81939	probable 50S ribos	1509	5	4.2	126	2	E90726	probable homeobox
1437	5	4.2	119	2	F81165	50S ribosomal prot	1510	5	4.2	126	2	PN0542	fortimicin KH epim
1438	5	4.2	119	2	S75418	ribosomal protein	1511	5	4.2	126	2	G72512	hypothetical prote
1439	5	4.2	119	2	B49905	protein secretion	1512	5	4.2	126	2	F85577	probable homeobox
1440	5	4.2	119	2	F60767	copia polypeptide	1513	5	4.2	126	2	A64811	ygs protein precu
1441	5	4.2	119	2	H84342	hypothetical prote	1514	5	4.2	126	2	AC1833	hypothetical prote
1442	5	4.2	119	2	E85882	hypothetical prote	1515	5	4.2	127	1	HSTR21	histone H2A, gonad
1443	5	4.2	119	2	B91038	hypothetical prote	1516	5	4.2	127	1	I53651	hydrophilic protei
1444	5	4.2	119	2	A65014	yfcC protein - Esc	1517	5	4.2	127	2	S50213	ubiquinol-cytochro
1445	5	4.2	119	2	AH0414	conserved hypotet	1518	5	4.2	127	2	S04574	Ig kappa chain pre
1446	5	4.2	119	2	S15218	vsdF protein - Sal	1519	5	4.2	127	2	B75353	hypothetical prote
1447	5	4.2	119	2	G93338	thioredoxin (trx-1	1520	5	4.2	127	2	F70717	hypothetical prote
1448	5	4.2	120	2	F72557	probable ribosomal	1521	5	4.2	127	2	T47929	hypothetical prote
1449	5	4.2	120	2	AG0808	conserved hypotet	1522	5	4.2	127	2	A72488	hypothetical prote
1450	5	4.2	120	2	AF0522	conserved hypotet	1523	5	4.2	127	2	AG2236	hypothetical prote
1451	5	4.2	120	2	B72647	hypothetical prote	1524	5	4.2	128	1	A57321	E48 antigen precu
1452	5	4.2	120	2	B42573	urfi20 - Paracoccu	1525	5	4.2	128	2	S31488	Ig kappa chain pre
1453	5	4.2	120	2	AE1408	conserved hypotet	1526	5	4.2	128	2	AH1332	large conductance
1454	5	4.2	120	2	F86434	protein F17F8.25 [1527	5	4.2	128	2	AH1703	large conductance
1455	5	4.2	120	2	E90491	conserved hypotet	1528	5	4.2	128	2	E72804	gp39 protein - Myc
1456	5	4.2	120	2	F96658	hypothetical prote	1529	5	4.2	128	2	AF0593	probable secreted
1457	5	4.2	120	2	F83043	hypothetical prote	1530	5	4.2	128	2	A75540	hypothetical prote
1458	5	4.2	120	2	AB1250	B. subtilis fmcA p	1531	5	4.2	128	2	B72459	hypothetical prote
1459	5	4.2	120	2	AF1612	B. subtilis fmcA p	1532	5	4.2	129	2	S45368	protein kinase C i
1460	5	4.2	121	2	B69149	ribosomal protein	1533	5	4.2	129	2	T49050	H+-transporting tw
1461	5	4.2	121	2	C39503	mercuric resistanc	1534	5	4.2	129	2	D81427	buforin I - Toad
1462	5	4.2	121	2	S70147	merD protein - Xan	1535	5	4.2	129	2	JC5397	cheY2 protein - Rh
1463	5	4.2	121	2	T18126	hypothetical prote	1536	5	4.2	129	2	S61838	root-specific prot
1464	5	4.2	121	2	S75563	hypothetical prote	1537	5	4.2	129	2	S28009	probable membrane
1465	5	4.2	121	2	T28128	hypothetical prote	1538	5	4.2	129	2	S69316	hypothetical prote
1466	5	4.2	121	2	I39846	small basic protei	1539	5	4.2	129	2	G75325	hypothetical prote
1467	5	4.2	121	2	AH0895	probable membrane	1540	5	4.2	129	2	H70536	hypothetical prote
1468	5	4.2	121	2	D65099	hypothetical 14.3	1541	5	4.2	129	2	C72678	hypothetical prote
1469	5	4.2	121	2	A91127	probable cytochrom	1542	5	4.2	129	2	A82508	conserved hypotet
1470	5	4.2	121	2	H55971	probable cytochrom	1543	5	4.2	129	2	G84192	hypothetical prote
1471	5	4.2	121	2	S34233	hypothetical prote	1544	5	4.2	129	2	C95050	conserved hypotet
1472	5	4.2	121	2	C82826	hypothetical prote	1545	5	4.2	129	2	H82866	hypothetical prote
1473	5	4.2	121	2	H86757	hypothetical prote	1546	5	4.2	129	2	G97920	conserved hypotet
1474	5	4.2	121	2	T31297	hypothetical prote	1547	5	4.2	129	2	A12799	conserved hypotet
1475	5	4.2	122	2	F70510	hypothetical prote	1548	5	4.2	129	2	C97579	heeb family protei
1476	5	4.2	122	2	S39457	pollen allergen Ph	1549	5	4.2	130	1	HSXLA1	histone H2A.1 - Af
1477	5	4.2	122	2	S72820	hypothetical prote	1550	5	4.2	130	1	F64376	hypothetical prote
1478	5	4.2	122	2	T50350	hypothetical prote	1551	5	4.2	130	2	I51445	histone H2A - Afri
1479	5	4.2	122	2	D95281	hypothetical prote	1552	5	4.2	130	2	T18510	hypothetical prote
1480	5	4.2	123	2	E75606	hypothetical prote	1553	5	4.2	130	2	T49245	cytoplasmic riboso
1481	5	4.2	123	2	A72520	hypothetical prote	1554	5	4.2	130	2	B86213	hypothetical prote
1482	5	4.2	123	2	G97885	hypothetical prote	1555	5	4.2	130	2	S20945	ribosomal protein
1483	5	4.2	124	1	HSURH2	histone H2A, embry	1556	5	4.2	130	2	A69216	translation initia
1484	5	4.2	124	1	HSUR7M	histone H2A, embry	1557	5	4.2	130	2	A75010	hypothetical prote
1485	5	4.2	124	1	HSURH9	histone H2A, embry	1558	5	4.2	130	2	T08532	traH protein - Ent
1486	5	4.2	124	2	A25077	histone H2A.2 - se	1559	5	4.2	130	2	S22993	potassium channel
1487	5	4.2	124	2	B86557	dnak suppressor [1	1560	5	4.2	130	2	A35633	hypothetical prote
1488	5	4.2	124	2	G72066	dnak suppressor -	1561	5	4.2	130	2	D84353	hypothetical prote
1489	5	4.2	124	2	A71518	probable dnak supp	1562	5	4.2	130	2	A87376	conserved hypotet

1563 5 4.2 130 2 G87422 hypothetical prote
1564 5 4.2 130 2 E90209 hypothetical prote
1565 5 4.2 130 2 AD1152 hypothetical prote
1566 5 4.2 130 2 AE1511 hypothetical prote
1567 5 4.2 130 2 F87384 cytochrome c famil
1568 5 4.2 130 2 A83340 hypothetical prote
1569 5 4.2 130 2 B72702 hypothetical prote
1570 5 4.2 130 2 A11875 hypothetical prote
1571 5 4.2 131 1 HSKLA2 histone H2A.2 - Af
1572 5 4.2 131 1 IEBC1 insB protein homol
1573 5 4.2 131 2 E69488 SSU ribosomal prot
1574 5 4.2 131 2 B72523 hypothetical prote
1575 5 4.2 131 2 T34900 probable secreted
1576 5 4.2 131 2 T46455 hypothetical prote
1577 5 4.2 131 2 F70892 hypothetical prote
1578 5 4.2 131 2 AF3339 precorrin-3b C17-m
1579 5 4.2 131 2 G72483 hypothetical prote
1580 5 4.2 131 2 A87355 hypothetical prote
1581 5 4.2 131 2 D87328 conserved hypothet
1582 5 4.2 131 2 D83173 hypothetical prote
1583 5 4.2 131 2 T34521 hypothetical prote
1584 5 4.2 131 2 AF3338 hypothetical prote
1585 5 4.2 131 2 F70920 probable moaB prot
1586 5 4.2 132 1 D39741 cytochrome c bioge
1587 5 4.2 132 1 A69340 conserved hypothet
1588 5 4.2 132 2 B89834 teichoic acid bios
1589 5 4.2 132 2 B81600 conserved hypothet
1590 5 4.2 132 2 B72307 conserved hypothet
1591 5 4.2 132 2 S69753 hypothetical prote
1592 5 4.2 132 2 B83299 hypothetical prote
1593 5 4.2 132 2 A13242 conserved hypothet
1594 5 4.2 132 2 D75548 hypothetical prote
1595 5 4.2 133 2 B71838 NADH2 dehydrogenas
1596 5 4.2 133 2 D64677 NADH2 dehydrogenas
1597 5 4.2 133 2 T40979 hypothetical HIT-f
1598 5 4.2 133 2 D69836 hypothetical prote
1599 5 4.2 133 2 A81119 conserved hypothet
1600 5 4.2 133 2 D97680 hypothetical prote
1601 5 4.2 133 2 G70651 hypothetical prote
1602 5 4.2 133 2 S38161 hypothetical prote
1603 5 4.2 134 1 P71142 whey acidic protei
1604 5 4.2 134 2 P71142 H+-transporting tw
1605 5 4.2 134 2 S73181 hydroxymethylbilan
1606 5 4.2 134 2 S12908 Ig kappa chain pre
1607 5 4.2 134 2 S11245 conserved hypothet
1608 5 4.2 134 2 A81676 hypothetical prote
1609 5 4.2 134 2 H70814 hypothetical prote
1610 5 4.2 134 2 E81133 hypothetical prote
1611 5 4.2 134 2 A26986 hypothetical prote
1612 5 4.2 134 2 T05417 hypothetical prote
1613 5 4.2 134 2 G84039 hypothetical prote
1614 5 4.2 134 2 AD2803 conserved hypothet
1615 5 4.2 134 2 D97582 hypothetical prote
1616 5 4.2 134 2 P70754 FtsW protein - Str
1617 5 4.2 135 2 G87686 hypothetical prote
1618 5 4.2 135 2 S31913 Ig gamma-2A chain
1619 5 4.2 135 2 B81097 transcription regu
1620 5 4.2 135 2 S58692 DNA-binding factor
1621 5 4.2 135 2 E75137 hypothetical prote
1622 5 4.2 135 2 H72471 hypothetical prote
1623 5 4.2 135 2 D82539 hypothetical prote
1624 5 4.2 135 2 H87680 hypothetical prote
1625 5 4.2 135 2 A12459 hypothetical prote
1626 5 4.2 135 2 G84212 hypothetical prote
1627 5 4.2 135 2 H70611 hypothetical prote
1628 5 4.2 135 2 T46448 ribosomal protein
1629 5 4.2 136 2 T00569 hypothetical prote
1630 5 4.2 136 2 C71143 hypothetical prote
1631 5 4.2 136 2 C85495 yacL protein - Esc
1632 5 4.2 136 2 G64734 hypothetical prote
1633 5 4.2 136 2 C90644 hypothetical prote
1634 5 4.2 136 2 T47771 probable transposa
1635 5 4.2 136 2 T35632

1636 5 4.2 136 2 T37117 probable transposa
1637 5 4.2 136 2 S46292 probable DNA mleme
1638 5 4.2 136 2 B82787 hypothetical prote
1639 5 4.2 136 2 B95333 hypothetical prote
1640 5 4.2 136 2 A10124 probable prepilin
1641 5 4.2 137 1 WVRT whey acidic protei
1642 5 4.2 137 2 D24773 protein-tyrosine k
1643 5 4.2 137 2 S73212 ribosomal protein
1644 5 4.2 137 2 T07997 ribosomal protein
1645 5 4.2 137 2 AG0029 mechanosensitive i
1646 5 4.2 137 2 F69046 conserved hypothet
1647 5 4.2 137 2 C72245 hypothetical prote
1648 5 4.2 137 2 H75615 hypothetical prote
1649 5 4.2 137 2 A83852 hypothetical prote
1650 5 4.2 137 2 G96779 probable RING zinc
1651 5 4.2 137 2 T33841 hypothetical prote
1652 5 4.2 137 2 H82537 TonB protein Xr260
1653 5 4.2 137 2 A11985 nitrogen fixation
1654 5 4.2 138 1 DMPG desmin - pig (frag
1655 5 4.2 138 1 Q088B2 Uu73 glycoprotein
1656 5 4.2 138 2 PQ0125 major merozoite su
1657 5 4.2 138 2 H75120 translation initia
1658 5 4.2 138 2 G71010 translation initia
1659 5 4.2 138 2 D71172 hypothetical prote
1660 5 4.2 138 2 C53291 probable B-ring cy
1661 5 4.2 138 2 S72933 hypothetical prote
1662 5 4.2 138 2 G82533 50S ribosomal prot
1663 5 4.2 138 2 S34725 regulation protein
1664 5 4.2 139 1 H64005 conserved hypothet
1665 5 4.2 139 2 S22417 3-oxoacyl-lacyl-ca
1666 5 4.2 139 2 B64324 ribosomal protein
1667 5 4.2 139 2 H81812 duplicated hypotthe
1668 5 4.2 139 2 H85077 hypothetical prote
1669 5 4.2 139 2 C70718 hypothetical prote
1670 5 4.2 139 2 T02564 hypothetical prote
1671 5 4.2 139 2 H84729 hypothetical prote
1672 5 4.2 139 2 C83526 hypothetical prote
1673 5 4.2 139 2 H84376 hypothetical prote
1674 5 4.2 139 2 C87544 hypothetical prote
1675 5 4.2 139 2 F95916 hypothetical expor
1676 5 4.2 139 2 F82863 conserved hypothet
1677 5 4.2 139 2 G72710 hypothetical prote
1678 5 4.2 139 2 AC1380 hypothetical prote
1679 5 4.2 139 2 AB2127 hypothetical prote
1680 5 4.2 140 2 D81970 H+-transporting tw
1681 5 4.2 140 2 F81024 ATP synthase F1, e
1682 5 4.2 140 2 AH0292 conserved hypothet
1683 5 4.2 140 2 AD1205 molybdopterin conv
1684 5 4.2 140 2 S27658 hypothetical prote
1685 5 4.2 140 2 D70831 probable mmp94 pro
1686 5 4.2 140 2 AB2828 hypothetical prote
1687 5 4.2 140 2 JC5515 hemoglobin alpha-D
1688 5 4.2 140 2 B48329 hypothetical prote
1689 5 4.2 140 2 E86967 probable transcrip
1690 5 4.2 141 2 C83007 probable transcrip
1691 5 4.2 141 2 C70572 hypothetical prote
1692 5 4.2 141 2 T35230 probable integral
1693 5 4.2 141 2 S34629 hypothetical prote
1694 5 4.2 141 2 S29168 hemoglobin alpha-D
1695 5 4.2 141 2 F95986 conserved hypothet
1696 5 4.2 142 1 HNMJBV hemagglutinin homo
1697 5 4.2 142 2 S40866 hypothetical prote
1698 5 4.2 142 2 H72600 hypothetical prote
1699 5 4.2 142 2 F86443 unknown protein [i
1700 5 4.2 142 2 B87334 hypothetical prote
1701 5 4.2 142 2 A60514 hemoglobin alpha-2
1702 5 4.2 142 2 S27050 hemoglobin alpha-2
1703 5 4.2 142 2 AB3401 hypothetical membr
1704 5 4.2 143 2 S07631 histone H2A.X - hu
1705 5 4.2 143 2 I48406 histone H2A.X - mo
1706 5 4.2 143 2 D70390 glycine cleavage s
1707 5 4.2 143 2 B69203 conserved hypothet
1708 5 4.2 143 2 D72741 hypothetical prote

1709	5	4.2	143	2	B72703	hypothetical prote	1782	5	4.2	148	2	JX0348	acrosomal major pr
1710	5	4.2	143	2	F84390	hypothetical prote	1783	5	4.2	148	2	JA0166	prolamin 14 precu
1711	5	4.2	143	2	G69826	transcription regu	1784	5	4.2	148	2	T02665	probable prolamin
1712	5	4.2	143	2	C75173	hypothetical prote	1785	5	4.2	148	2	B90025	molybdopterin conv
1713	5	4.2	143	2	B71159	hypothetical prote	1786	5	4.2	148	2	AD3489	cytochrome o ubiqu
1714	5	4.2	143	2	E71041	hypothetical prote	1787	5	4.2	148	2	T35069	probable membrane
1715	5	4.2	144	1	XSSMA	subtilisin inhibit	1788	5	4.2	148	2	E86741	hypothetical prote
1716	5	4.2	144	2	S78178	succinate dehydrog	1789	5	4.2	148	2	T33937	hypothetical prote
1717	5	4.2	144	2	JN0480	phospholipase A2 (1790	5	4.2	148	2	I39576	ncck protein - Aic
1718	5	4.2	144	2	S11244	Ig gamma-2a chain	1791	5	4.2	148	2	T31516	hypothetical prote
1719	5	4.2	144	2	T03563	probable ribosomal	1792	5	4.2	149	1	FWRZP7	prolamin 7 precurs
1720	5	4.2	144	2	G82092	conserved hypothet	1793	5	4.2	149	2	G72250	ribosomal protein
1721	5	4.2	144	2	AH2458	hypothetical prote	1794	5	4.2	149	2	JH0219	prolamin precursor
1722	5	4.2	144	2	A63399	hypothetical prote	1795	5	4.2	149	2	A13293	4-hydroxybenzoyl-C
1723	5	4.2	144	2	A82263	hypothetical prote	1796	5	4.2	149	2	T31970	hypothetical prote
1724	5	4.2	144	2	D83152	hypothetical prote	1797	5	4.2	149	2	S27692	tcuJ protein - Str
1725	5	4.2	144	2	A87699	hypothetical prote	1798	5	4.2	149	2	B86579	CF667 hypothetical
1726	5	4.2	144	2	D70841	hypothetical prote	1799	5	4.2	149	2	E72046	conserved hypothet
1727	5	4.2	144	2	D70949	hypothetical prote	1800	5	4.2	149	2	H72113	hypothetical prote
1728	5	4.2	144	2	F70895	probable PE protei	1801	5	4.2	149	2	C86509	hypothetical prote
1729	5	4.2	144	2	T01823	hypothetical prote	1802	5	4.2	149	2	F75327	hypothetical prote
1730	5	4.2	144	2	T50622	hypothetical prote	1803	5	4.2	149	2	JN0436	microbial serine p
1731	5	4.2	144	2	AH1997	hypothetical prote	1804	5	4.2	150	2	F53312	triose-phosphate i
1732	5	4.2	145	1	PSBOA	phospholipase A2 (1805	5	4.2	150	2	S54847	succinate-CoA liga
1733	5	4.2	145	2	S65309	probable membrane	1806	5	4.2	150	2	S18394	troponin C isoform
1734	5	4.2	145	2	S77196	hypothetical prote	1807	5	4.2	150	2	JH0220	prolamin precursor
1735	5	4.2	145	2	AH2102	hypothetical prote	1808	5	4.2	150	2	JA0168	prolamin 4a precu
1736	5	4.2	145	2	D70946	hypothetical prote	1809	5	4.2	150	2	G83581	probable transcrip
1737	5	4.2	145	2	S23005	tram protein - Bac	1810	5	4.2	150	2	T19659	hypothetical prote
1738	5	4.2	145	2	H75108	hypothetical prote	1811	5	4.2	150	2	B75528	hypothetical prote
1740	5	4.2	145	2	A29220	cobalamin biosynth	1812	5	4.2	150	2	C84041	hypothetical prote
1741	5	4.2	145	2	E97694	hypothetical prote	1813	5	4.2	150	2	D87652	hypothetical prote
1742	5	4.2	145	2	D90818	hypothetical prote	1814	5	4.2	150	2	A55883	actin-filament-asa
1743	5	4.2	145	2	H85677	unknown protein en	1815	5	4.2	150	2	A97536	hypothetical prote
1744	5	4.2	145	2	AG2617	conserved hypothet	1816	5	4.2	151	2	J50672	prolamin 13K precu
1745	5	4.2	145	2	T33665	hypothetical prote	1817	5	4.2	151	2	S47056	hypothetical prote
1746	5	4.2	146	1	PSDG	phospholipase A2 (1818	5	4.2	151	2	F87698	patch repair prote
1747	5	4.2	146	1	S34049	phospholipase A2 (1819	5	4.2	151	2	T28645	transcription fact
1748	5	4.2	146	1	HBAK	hemoglobin beta ch	1820	5	4.2	151	2	H69991	hypothetical prote
1749	5	4.2	146	1	C75189	hemoglobin beta ch	1821	5	4.2	151	2	AD0037	teillurite resistan
1750	5	4.2	146	2	H81445	hypothetical prote	1822	5	4.2	151	2	H83041	conserved hypothet
1751	5	4.2	146	2	H81977	(3R)-hydroxymyrist	1823	5	4.2	151	2	C72555	hypothetical prote
1752	5	4.2	146	2	A64866	probable acetyltra	1824	5	4.2	152	1	VGBEY9	early glycoprotein
1753	5	4.2	146	2	T10511	hypothetical prote	1825	5	4.2	152	2	AB2327	hypothetical prote
1754	5	4.2	146	2	T08537	tram protein - Ent	1826	5	4.2	152	2	A70800	hypothetical prote
1755	5	4.2	146	2	G75544	hypothetical prote	1827	5	4.2	152	2	F83224	conserved hypothet
1756	5	4.2	146	2	T08344	hypothetical prote	1828	5	4.2	152	2	S27150	gene 30.3 protein
1757	5	4.2	146	2	C96708	hypothetical prote	1829	5	4.2	152	2	C72662	hypothetical prote
1758	5	4.2	146	2	A87513	hypothetical prote	1830	5	4.2	153	1	XKPOC1	proteinase inhibit
1759	5	4.2	146	2	H87622	hypothetical prote	1831	5	4.2	153	1	B69215	conserved hypothet
1760	5	4.2	146	2	JC4902	ionized calcium bi	1832	5	4.2	153	2	A34132	vasotocin / neurop
1761	5	4.2	146	2	T17559	glycine-rich prote	1833	5	4.2	153	2	A28543	peptidoglycan-asso
1762	5	4.2	146	2	E70500	hypothetical prote	1834	5	4.2	153	2	T27157	hypothetical prote
1763	5	4.2	146	2	S57956	ovine vascular end	1835	5	4.2	153	2	H85638	hypothetical prote
1764	5	4.2	146	2	S03290	glucoprotein, 190K	1836	5	4.2	154	2	JN0410	myoglobin - Europe
1765	5	4.2	147	2	D86795	glutathione peroxi	1837	5	4.2	154	2	AD2388	hypothetical prote
1766	5	4.2	147	2	B97356	ribosomal protein	1838	5	4.2	154	2	B75358	hypothetical prote
1767	5	4.2	147	2	F83505	probable peptide d	1839	5	4.2	154	2	H83160	hypothetical prote
1768	5	4.2	147	2	G83301	conserved hypothet	1840	5	4.2	154	2	D97690	hypothetical prote
1769	5	4.2	147	2	H70514	hypothetical prote	1841	5	4.2	154	2	A12915	conserved hypothet
1770	5	4.2	147	2	C47056	cnr regulatory pro	1842	5	4.2	154	2	AG3461	hypothetical prote
1771	5	4.2	147	2	S20527	hypothetical prote	1843	5	4.2	154	2	A85431	hypothetical prote
1772	5	4.2	147	2	C31844	kilB protein - Str	1844	5	4.2	154	2	D83516	hypothetical prote
1773	5	4.2	147	2	S76645	hypothetical prote	1845	5	4.2	154	2	F72741	hypothetical prote
1774	5	4.2	147	2	D83302	hypothetical prote	1846	5	4.2	155	2	S17912	hypothetical prote
1775	5	4.2	147	2	I55617	allograft inflama	1847	5	4.2	155	2	S59956	NADH2 dehydrogenas
1776	5	4.2	147	2	B82270	probable fibrinola	1848	5	4.2	155	2	S48414	probable membrane
1777	5	4.2	147	2	E75003	hypothetical prote	1849	5	4.2	155	2	A30048	hypothetical prote
1778	5	4.2	148	1	JC4649	pseudoazurin precu	1850	5	4.2	155	2	D87399	cytochrome c famil
1779	5	4.2	148	1	VGBE48	early glycoprotein	1851	5	4.2	155	2	A10545	hypothetical prote
1780	5	4.2	148	2	S25593	60S ribosomal prot	1852	5	4.2	155	2	F84214	hypothetical prote
1781	5	4.2	148	2	S60001	60S ribosomal prot	1853	5	4.2	156	2	T33725	ribonuclease H (BC
							1854	5	4.2	156	2	G90434	hypothetical prote

1855 5 4.2 156 2 C97503
 1856 5 4.2 156 2 A87567
 1857 5 4.2 156 2 B84230
 1858 5 4.2 156 2 D72258
 1859 5 4.2 157 2 S41891
 1860 5 4.2 157 2 A55046
 1861 5 4.2 157 2 B30177
 1862 5 4.2 157 2 B72606
 1863 5 4.2 157 2 A72460
 1864 5 4.2 157 2 T05818
 1865 5 4.2 157 2 T34552
 1866 5 4.2 157 2 D82741
 1867 5 4.2 158 2 A81555
 1868 5 4.2 158 2 A82256
 1869 5 4.2 158 2 F72725
 1870 5 4.2 158 2 G70323
 1871 5 4.2 158 2 T08311
 1872 5 4.2 158 2 C8314
 1873 5 4.2 158 2 H8316
 1874 5 4.2 158 2 H83405
 1875 5 4.2 158 2 T26444
 1876 5 4.2 158 2 J05297
 1877 5 4.2 158 2 A12718
 1878 5 4.2 158 2 E97500
 1879 5 4.2 159 2 A61197
 1880 5 4.2 159 2 S42986
 1881 5 4.2 159 2 I50810
 1882 5 4.2 159 2 T02076
 1883 5 4.2 159 2 A30125
 1884 5 4.2 159 2 A83654
 1885 5 4.2 159 2 B86518
 1886 5 4.2 159 2 A72104
 1887 5 4.2 159 2 E72731
 1888 5 4.2 159 2 H75391
 1889 5 4.2 159 2 P64394
 1890 5 4.2 159 2 B70635
 1891 5 4.2 159 2 B72764
 1892 5 4.2 159 2 S61040
 1893 5 4.2 159 2 A81899
 1894 5 4.2 160 2 A40060
 1895 5 4.2 160 2 D4983
 1896 5 4.2 160 2 T50732
 1897 5 4.2 160 2 E35882
 1898 5 4.2 161 2 A45000
 1899 5 4.2 161 2 S73630
 1900 5 4.2 161 2 F87277
 1901 5 4.2 161 2 E75043
 1902 5 4.2 161 2 H87034
 1903 5 4.2 161 2 G87048
 1904 5 4.2 161 2 E87455
 1905 5 4.2 162 2 T09638
 1906 5 4.2 162 2 A43428
 1907 5 4.2 162 2 F84380
 1908 5 4.2 162 2 G64351
 1909 5 4.2 162 2 T51723
 1910 5 4.2 162 2 D30502
 1911 5 4.2 162 2 A80915
 1912 5 4.2 162 2 E70195
 1913 5 4.2 162 2 B83423
 1914 5 4.2 162 2 F82084
 1915 5 4.2 162 2 H70927
 1916 5 4.2 162 2 B31225
 1917 5 4.2 162 2 H86071
 1918 5 4.2 162 2 S70372
 1919 5 4.2 162 2 A81950
 1920 5 4.2 163 1 Q08B81
 1921 5 4.2 163 2 T21418
 1922 5 4.2 163 2 A83002
 1923 5 4.2 163 2 A13322
 1924 5 4.2 163 2 A81694
 1925 5 4.2 163 2 A81210
 1926 5 4.2 163 2 JF0022
 1927 5 4.2 163 2 T38267

1928 5 4.2 163 2 D75504
 1929 5 4.2 163 2 D87293
 1930 5 4.2 163 2 A70847
 1931 5 4.2 163 2 AC3197
 1932 5 4.2 164 1 G65085
 1933 5 4.2 164 2 AB1900
 1934 5 4.2 164 2 S27650
 1935 5 4.2 164 2 F91113
 1936 5 4.2 164 2 F85958
 1937 5 4.2 164 2 AB0885
 1938 5 4.2 164 2 I39548
 1939 5 4.2 164 2 S20525
 1940 5 4.2 164 2 A83262
 1941 5 4.2 164 2 AE1076
 1942 5 4.2 164 2 G87705
 1943 5 4.2 164 2 H85044
 1944 5 4.2 164 2 T35225
 1945 5 4.2 165 2 E90069
 1946 5 4.2 165 2 H70635
 1947 5 4.2 165 2 E83328
 1948 5 4.2 165 2 AH3359
 1949 5 4.2 165 2 T32451
 1950 5 4.2 165 2 E70385
 1951 5 4.2 165 2 B72645
 1952 5 4.2 165 2 G95074
 1953 5 4.2 165 2 D75292
 1954 5 4.2 166 2 AH3508
 1955 5 4.2 166 2 H98280
 1956 5 4.2 166 2 I53166
 1957 5 4.2 166 2 G81235
 1958 5 4.2 166 2 S09223
 1959 5 4.2 166 2 S09224
 1960 5 4.2 166 2 S09225
 1961 5 4.2 166 2 S40144
 1962 5 4.2 166 2 T13128
 1963 5 4.2 166 2 S13219
 1964 5 4.2 166 2 AC0120
 1965 5 4.2 166 2 AG0178
 1966 5 4.2 166 2 F86767
 1967 5 4.2 166 2 H89044
 1968 5 4.2 166 2 A40783
 1969 5 4.2 167 1 LTHU
 1970 5 4.2 167 2 E86337
 1971 5 4.2 167 2 S33707
 1972 5 4.2 167 2 T44271
 1973 5 4.2 167 2 G82077
 1974 5 4.2 167 2 AG0285
 1975 5 4.2 167 2 D75510
 1976 5 4.2 167 2 S49628
 1977 5 4.2 167 2 D82448
 1978 5 4.2 167 2 A83493
 1979 5 4.2 167 2 E95875
 1980 5 4.2 168 1 CUSP
 1981 5 4.2 168 2 S26079
 1982 5 4.2 168 2 S77193
 1983 5 4.2 168 2 D75140
 1984 5 4.2 168 2 T30315
 1985 5 4.2 168 2 G83068
 1986 5 4.2 168 2 H87611
 1987 5 4.2 168 2 AC3212
 1988 5 4.2 168 2 E83437
 1989 5 4.2 168 2 A83350
 1990 5 4.2 168 2 S52994
 1991 5 4.2 168 2 A25522
 1992 5 4.2 169 1 S20501
 1993 5 4.2 169 2 T04207
 1994 5 4.2 169 2 D87294
 1995 5 4.2 169 2 B48489
 1996 5 4.2 169 2 F70799
 1997 5 4.2 169 2 D72765
 1998 5 4.2 169 2 S17940
 1999 5 4.2 169 2 T50907
 2000 5 4.2 169 2 A12436

conserved hypothet
 Fur family protein
 hypothetical prote
 hypothetical prote
 hydrogenase 2 matu
 NADH dehydrogenase
 2,3,4,5-tetrahydro
 hydrogenase 2 matu
 probable processin
 hydrogenase-2 comp
 aut protein - Alca
 hypothetical prote
 hypothetical prote
 probable secreted
 hypothetical prote
 hypothetical prote
 probable nitrate r
 hypothetical prote
 thioredoxin peroxi
 thiol peroxidase P
 probable competenc
 hypothetical prote
 conserved hypothet
 hypothetical prote
 PTS system IIA com
 conserved hypothet
 fxa protein [mpo
 shikimate kinase
 lepton precursor -
 50S ribosomal prot
 membrane protein -
 membrane protein -
 membrane protein -
 premembrane protei
 protein gp42 - pha
 s18 protein - frui
 conserved hypothet
 conserved hypothet
 hypothetical prote
 protein B0238.12
 18.2K protein - ph
 lepton precursor -
 plastocyanin [simi
 DNA-damage repair
 glutathione peroxi
 fibrial protein V
 thiol peroxidase
 hypothetical prote
 FUN81 protein - ye
 hypothetical prote
 hypothetical prote
 hypothetical prote
 plastocyanin precu
 micrococcal nuclea
 general secretion
 hypothetical prote
 chpC protein - Pse
 hypothetical prote
 transcription regu
 hypothetical prote
 hypothetical prote
 probable sigma-70
 arabinogalactan-li
 major merozoite su
 probable glutathio
 phospholipid-hydro
 Atp synthase F0, B
 nitrate reductase
 hypothetical prote
 hypothetical prote
 export protein xps
 2-vinyl bacterioch
 hypothetical prote

ALIGNMENTS

RESULT 1
PC1306
genome polyprotein NS4a epitope containing region (isolate BR36-20) - hepatitis C virus
C:Species: hepatitis C virus
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C:Accession: PC1306
R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.
Biochem. Biophys. Res. Commun. 192, 635-641, 1993
A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.
A:Reference number: PC1300; MUID:93249436; PMID:7683463
A:Accession: PC1306
A:Molecule type: mRNA
A:Residues: 1-209 <STU>
A:Cross-references: UNIPROT:Q81594; UNIPARC:UPI0000033B46; DDBJ:D14600; NID:G303584; PID
A:Experimental source: blood
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: nonstructural protein; polyprotein

Query Match 70.3%; Score 83; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.7e-73;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGGKPAIVPDKEVLYQQYD 60
|||||
DB 92 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGGKPAIVPDKEVLYQQYD 151
|||||

QY 61 EMRECSQAAPYIEQAQVIAHQPK 83
|||||
DB 152 EMRECSQAAPYIEQAQVIAHQPK 174
|||||

RESULT 2
PC1307
genome polyprotein NS4a epitope containing region (isolate HD10-1) - hepatitis C virus
C:Species: hepatitis C virus
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 31-Dec-2004
C:Accession: PC1307
R:Stuyver, L.; Van Arnhem, W.; Wyseur, A.; Deleys, R.; Maertens, G.
Biochem. Biophys. Res. Commun. 192, 635-641, 1993
A:Title: Analysis of the putative E1 envelope and NS4a epitope regions of HCV type 3.
A:Reference number: PC1300; MUID:93249436; PMID:7683463
A:Accession: PC1307
A:Molecule type: mRNA
A:Residues: 1-142 <STU>
A:Cross-references: UNIPROT:Q68870; UNIPARC:UPI0000178536; DDBJ:D14602
A:Experimental source: blood
C:Keywords: polyprotein

Query Match 39.8%; Score 47; DB 2; Length 142;
Best Local Similarity 100.0%; Pred. No. 2e-38;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGGKPA 47
|||||
DB 25 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGGKPA 71
|||||

RESULT 3
PS0326
polyprotein - hepatitis C virus (isolate Fla) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C:Accession: PS0326
R:Li, J.S.; Tong, S.P.; Vitvitski, L.; Lepot, D.; Trepo, C.
Gene 105, 167-172, 1991
A:Title: Two French genotypes of hepatitis C virus: homology of the predominant genotype
A:Reference number: PS0326; MUID:92039028; PMID:1718820
A:Accession: PS0326

A:Molecule type: genomic RNA
A:Residues: 1-492 <LIJ>
A:Cross-references: UNIPROT:Q91PES; UNIPROT:Q36579; UNIPROT:Q36610; UNIPROT:Q03463; UNIP
PARC:UPI0000178532; GB:M60220
A:Note: this sequence corresponds to nonstructural protein NS3 region
A:Note: translation of the nucleotide sequence is not complete
C:Keywords: polyprotein

Query Match 11.0%; Score 13; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 9.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
|||||
DB 215 GGVLAALAAAYCLIS 227
|||||

RESULT 4
JQ1366
polyprotein - hepatitis C virus (French isolate) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 31-Dec-2004
C:Accession: JQ1366
R:Krensdorff, D.; Porchon, C.; Kim, J.P.; Reyes, G.R.; Brechot, C.
J. Gen. Virol. 72, 2557-2561, 1991
A:Title: Partial nucleotide sequence analysis of a French hepatitis C virus: implication
A:Reference number: JQ1366; MUID:92013977; PMID:1655961
A:Accession: JQ1366
A:Molecule type: genomic RNA
A:Residues: 1-716 <KRS>
A:Cross-references: UNIPROT:Q9PX22; UNIPARC:UPI0000178531
C:Keywords: glycoprotein; polyprotein
F:84,90,97,115,143,199,223,243,290,312/Binding site: carbohydrate (Asn) (covalent) #stat

Query Match 11.0%; Score 13; DB 2; Length 716;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
|||||
DB 613 GGVLAALAAAYCLIS 625
|||||

RESULT 5
GNWVC3
genome polyprotein - hepatitis C virus (strain HCV-1)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain HCV-1)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 09-Jul-2004
C:Accession: A39166; PQ0403; PQ0404
R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coi
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A:Title: Genetic organization and diversity of the hepatitis C virus.
A:Reference number: A39166; MUID:91172826; PMID:1848704
A:Accession: A39166
A:Molecule type: mRNA
A:Residues: 1-3011 <CHO>
A:Cross-references: UNIPROT:P26664; UNIPARC:UPI0000131E19; GB:M62321; NID:G329873; PIDN:
R:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I
J. Gen. Virol. 73, 1131-1141, 1992
A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e
A:Reference number: PQ0393; MUID:92268871; PMID:1316939
A:Accession: PQ0403
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CHA>
A:Cross-references: UNIPARC:UPI0000174A00; DDBJ:D10128
A:Experimental source: isolates E-b16
A:Accession: PQ0404
A:Status: preliminary
A:Molecule type: genomic RNA
A:Residues: 1577-1633 <CH2>
A:Cross-references: UNIPARC:UPI0000174A00

```
A:Experimental source: isolates E-bl7
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypeptide
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,2240,23
Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLIS 30
Db 1664 GGVLAALAAAYCLIS 1676

RESULT 6
GNVUCH
genome polypeptide - hepatitis C virus (strain H)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polypeptide
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: A36814; A41546
R:Inchausti, G.; Zebadee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A:Description: Genomic structure of the human prototype strain H of hepatitis C virus: C
A:Reference number: A36814
A:Accession: A36814
A:Molecule type: genomic RNA
A:Residues: 1-3011 <INC>
A:Cross-references: UNIPROT:P27958; UNIPARC:UPI000004ACD7; GB:M67463; NID:9329737; PIDN:
R:Inchausti, G.; Zebadee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: compari
A:Reference number: A41546; MUID:92052256; PMID:1658800
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural
F:1-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypeptide
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23
Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLIS 30
Db 1664 GGVLAALAAAYCLIS 1676
```

RESULT 7

```
S40770
genome polypeptide - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polypeptide
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S40770; PC1285
R:Okamoto, H.
submitted to the EMBL Data Library, March 1992
A:Reference number: S40770
A:Accession: S40770
A:Molecule type: genomic RNA
A:Residues: 1-3011 <OKA>
A:Cross-references: UNIPROT:Q03463; UNIPARC:UPI00000F0182; EMBL:D10749; NID:G221586; PID
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tada,
Jpn. J. Exp. Med. 60, 167-177, 1990
A:Title: The 5'-terminal sequence of the hepatitis C virus genome.
A:Reference number: PC1284; MUID:91013116; PMID:2170712
A:Accession: PC1285
A:Molecule type: genomic RNA
A:Residues: 1-513 <OK2>
A:Cross-references: UNIPARC:UPI00000F4078; GB:D00831; NID:G221511; PIDN:BAA00705.1; PID:
A:Experimental source: isolate HC-J1
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polypeptide
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <NS4a>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <NS4b>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
Query Match 11.0%; Score 13; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.00043;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCLIS 30
Db 1664 GGVLAALAAAYCLIS 1676

RESULT 8
A45573
genome polypeptide - hepatitis C virus (strain JT)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus genome polypeptide
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: A45573
R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,
Virus Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: s
A:Reference number: A45573; MUID:92295714; PMID:1318627
A:Accession: A45573
A:Molecule type: DNA
A:Residues: 1-3010 <YAN>
A:Cross-references: UNIPROT:Q00269; UNIPARC:UPI0000131E29; GB:D01168; GB:D01171; NID:922
A:Experimental source: HCV-JT
A:Note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:106207)
C:Superfamily: hepatitis C virus genome polypeptide
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
```


Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 12

S18030
genome polyprotein - hepatitis C virus (isolate JKI)
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (isolate JKI) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Variety: isolate JKI
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: S18030; S33570; R48332; S18029
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A;Description: A whole genome of hepatitis C virus cDNA was isolated from a single patient
A;Reference number: S18028
A;Accession: S18030
A;Molecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: UNIPROT:Q68949; UNIPARC:UPI00000F2A81; EMBL:X61596; NID:G59478; PIDN
A;Experimental source: isolate JKI from an individual
R;Honda, M.; Kaneko, S.; Uchida, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A;Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A;Reference number: A48332; MUID:93119270; PMID:8380322
A;Accession: S33570
A;Molecule type: genomic RNA
A;Residues: 1-547, T', 549-621, V', 623-624, S', 626-652, DL', 655-761, T', 763-782 <HOW>
A;Cross-references: UNIPARC:UPI00001749FF; EMBL:X61591
A;Note: this sequence is inconsistent with the nucleotide translation
A;Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320 as Trp, and TTC for residue 771 as Ser
A;Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBI:P:121748)
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MBE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As

Query Match 10.2%; Score 12; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 13

B84219
4-hydroxybenzoate octaprenyltransferase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: B84219
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitthaus, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablonski, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Lid

A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84219
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <STO>
A;Cross-references: UNIPROT:Q9HRP0; UNIPARC:UPI00000636A1; GB:AE004437; NID:G10580202; P
C;Genetics: hhoA
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0279

Query Match 7.6%; Score 9; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAA 26
|||||
Db 161 GGVLAALAA 169

RESULT 14

PC2219
polypeptide - hepatitis C virus (type 5a) (fragments)
N;Contains: core protein; E1 (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A prote
C;Species: hepatitis C virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 31-Dec-2004
C;Accession: PC2219
R;Stuyver, L.; Arnhen, W.V.; Wyseur, A.; Maertens, G.
Biochem. Biophys. Res. Commun. 202, 1308-1314, 1994
A;Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the h
A;Reference number: PC2219; MUID:94338342; PMID:7520237
A;Accession: PC2219
A;Molecule type: mRNA
A;Residues: 1-876 <STU>
A;Cross-references: UNIPROT:Q81242; UNIPARC:UPI00001784FF; GB:L29577; GB:L29578; GB:L295
A;Experimental source: serum
C;Keywords: glycoprotein
F:1-191/Product: core #status predicted <COR>
F:68-78/Region: variable
F:192-247/Product: E1 (carboxyl end) #status predicted <ERE>
F:248-411/Product: E2/NS1 (amino end) #status predicted <ENR>
F:248-338/Region: E2
F:339-411/Region: NS1 (amino end)
F:412-783/Product: NS3 #status predicted <NSR>
F:784-837/Product: NS4A #status predicted <NSA>
F:838-876/Product: NS4B #status predicted <NSB>
F:281,287,294,312,340/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.6%; Score 9; DB 2; Length 876;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
|||||
Db 773 ACMSADLEV 781

RESULT 15

JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (isolate EUH1480) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A;Reference number: JC5620; MUID:97366593; PMID:9223423
A;Accession: JC5620
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: UNIPROT:O39928; UNIPARC:UPI0000174A01; GB:Y13184

A:Experimental source: genotype 5a, which predominates in South Africa
 A:Note: the translation of the nucleotide sequence is not complete in this paper
 C:Superfamily: hepatitis C virus genome polyprotein
 C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
 F:2-115/Product: capsid protein C #status predicted <CPC>
 F:116-191/Product: envelope protein M #status predicted <EPM>
 F:192-389/Product: major envelope protein E #status predicted <MEB>
 F:384-408/Region: hypervariable #status predicted
 F:390-730/Product: nonstructural protein NS1 #status predicted <NS1>
 F:731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
 F:1008-1616/Product: hepatitis virus #status predicted <NS3>
 F:1231-1238/Region: nucleotide-binding motif A (P-loop)
 F:1313-1318/Region: nucleotide-binding motif B
 F:1317-1320/Region: DEXH motif
 F:1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
 F:1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
 F:2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
 F:2210-2249/Region: interferon sensitivity determining #status predicted

Query Match 7.6%; Score 9; DB 1; Length 3014;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEV 9
 Db 1648 ACMSADLEV 1656
 |||||

RESULT 16
 S61982
 Probable membrane protein YOI002c - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein O2501; hypothetical protein UND327
 C:Species: Saccharomyces cerevisiae
 C:Date: 10-Apr-1996 #sequence revision 19-Apr-1996 #text_change 31-Dec-2004
 C:Accession: S61982; S66683; S72131
 R:Starky, F.; Uhlen, M.
 submitted to the EMBL Data Library, December 1995
 A:Reference number: S61981
 A:Accession: S61982
 A:Molecule type: DNA
 A:Residues: 1-327 <STB>
 A:Cross-references: UNIPROT:Q12442; UNIPARC:UPI000006AC11; EMBL:U43491; NID:g1150992; PI
 R:Petersson, B.; Starky, F.; Uhlen, M.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66682
 A:Accession: S66683
 A:Molecule type: DNA
 A:Residues: 1-327 <PET>
 A:Cross-references: UNIPARC:UPI000006AC11; EMBL:274744; NID:g1419762; PIDN:CAA99001.1; F
 R:Starky, F.; Holmberg, A.; Petersson, B.; Uhlen, M.
 Yeast 12, 1091-1095, 1996
 A:Title: The sequence of a 30 kb fragment on the left arm of chromosome XV from Sacchar
 A:Reference number: S72130; MUID:97051599; PMID:8896276
 A:Accession: S72131
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-327 <STW>
 A:Cross-references: UNIPARC:UPI000006AC11; EMBL:U43491; NID:g1150992; PIDN:AAC49478.1; F
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
 C:Genetics:
 A:Cross-references: SGD:S0005362
 A:Map position: 15L
 A:Note: YOI002c
 C:Superfamily: Adiponectin receptor protein and homologs
 C:Keywords: transmembrane protein
 F:93-109/Domain: transmembrane #status predicted <TM1>
 F:126-142/Domain: transmembrane #status predicted <TM2>
 F:164-180/Domain: transmembrane #status predicted <TM3>
 F:191-207/Domain: transmembrane #status predicted <TM4>
 F:225-241/Domain: transmembrane #status predicted <TM5>
 F:255-271/Domain: transmembrane #status predicted <TM6>
 F:294-310/Domain: transmembrane #status predicted <TM7>

Query Match 6.8%; Score 8; DB 2; Length 327;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 WYLLGGVL 21
 Db 257 WYLLGGVL 264
 |||||

RESULT 17
 AD3450
 proline racemase (EC 5.1.1.4) [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 R:DelVecchio, V.G.; Kapur, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
 .; Mazur, M.; Goltzman, E.; Selkov, E.; Eizer, P.H.; Hagiub, S.; O'Callaghan, D.; Letes
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AD3450
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-342 <KUR>
 A:Cross-references: UNIPROT:Q8YFD6; UNIPARC:UPI0000058118; GB:AE008917; PIDN:AAL52767.1;
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BME11586
 A:Map position: 1
 C:Superfamily: proline racemase
 C:Keywords: isomerase

Query Match 6.8%; Score 8; DB 2; Length 342;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ELGGKPAI 48
 Db 297 ELGGKPAI 304
 |||||

RESULT 18
 EG9889
 spore germination protein homolog yndE - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: EG9889
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, B.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Gall
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetel
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
 akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpatra, P.; Tognoni, A.; Tosato, V.; Uchiyana
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumstein, B.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: EG9889
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-363 <KUN>
 A:Cross-references: UNIPROT:O31809; UNIPARC:UPI0000060470; GB:Z99113; GB:AL009126; NID:9
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yndE
 C:Superfamily: spore germination protein

Query Match 6.8%; Score 8; DB 1; Length 363;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 8; Conservative 0; Mismatches 0; Gaps 0;

Qy 15 VLLGGVLA 22
Db 46 VLLGGVLA 53
|||||

RESULT 19
E70598
Probable PPE protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70598
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID:98295987; PMID:9634230
A: Accession: E70598
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-399 <COL>
A: Cross-references: UNIPROT:O05452; UNIPARC:UPI00000D4F1E; GB:Z94121; GB:AL123456; NID:G
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: PPE

Query Match 6.8%; Score 8; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALA 25
Db 160 GGVLAALA 167
|||||

RESULT 20
C84201
multidrug resistance protein homolog [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84201
R: Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A: Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omar, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A: Title: Genome sequence of Halobacterium species NRC-1.
A: Reference number: A84160; MUID:20504483; PMID:11016950
A: Accession: C84201
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-420 <STO>
A: Cross-references: UNIPROT:Q9HS33; UNIPARC:UPI000006361C; GB:AE004437; NID:gl0580037; E
C: Genetics:
A: Gene: yfmO2

Query Match 6.8%; Score 8; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALA 25
Db 389 GGVLAALA 396
|||||

RESULT 21
S54909

GTP cyclohydrolase I (EC 3.5.4.16) - Euglena gracilis (fragment)
C:Species: Euglena gracilis
C:Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C:Accession: S54909
R: Maier, J.; Witter, K.; Guetlich, M.; Ziegler, I.; Werner, T.; Ninnemann, H.
submitted to the EMBL Data Library, May 1995
A: Description: Homology cloning of GTP cyclohydrolase I from various unreleased eukaryot
A: Reference number: S54845
A: Accession: S54909
A: Status: preliminary
A: Molecule type: mRNA
A: Residues: 1-80 <MAI>
A: Cross-references: UNIPROT:P51597; UNIPARC:UPI000012B2A3; EMBL:Z49757; NID:g840757; PID
C: Superfamily: GTP cyclohydrolase I
C: Keywords: folate biosynthesis; GTP; hydrolase

Query Match 5.9%; Score 7; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLGL 89
Db 22 KGKVLGL 28
|||||

RESULT 22
H83065
hypothetical protein PA4642 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83065
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lazbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A: Reference number: A82950; MUID:20437337; PMID:10984043
A: Accession: H83065
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-96 <STO>
A: Cross-references: UNIPROT:Q9HVFO; UNIPARC:UPI00000C5DA4; GB:AE004878; GB:AE004091; NID
A: Experimental source: strain PA01
C: Genetics:
A: Gene: PA4642

Query Match 5.9%; Score 7; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 QAAPYIE 73
Db 82 QAAPYIE 88
|||||

RESULT 23
G83545
hypothetical protein PA0802 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: G83545
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lazbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A: Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A: Reference number: A82950; MUID:20437337; PMID:10984043
A: Accession: G83545
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-107 <STO>
A: Cross-references: UNIPROT:Q9I5D6; UNIPARC:UPI00000C515P; GB:AE004515; GB:AE004091; NID
A: Experimental source: strain PA01

C:Genetics:
A:Gene: PA0802

Query Match 5.9%; Score 7; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LIGGVLA 22
|||||
DB 98 LIGGVLA 104

RESULT 24

S35629 hypothetical protein - hepatitis C virus

C:Species: hepatitis C virus

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S35629

R:Sasahina, T.; Sakurai, T.; Watanabe, Y.; Kashima, K.; Suzuki, T.; Chiba, J.; Kita, Y.

Nucleic Acids Res. 21, 1037, 1993

A:Title: Nucleotide sequence of the hepatitis C virus genome from a patient negative for

A:Reference number: S35629; MUID:93197128; PMID:8383835

A:Accession: S35629

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-125 <SAR>

A:Cross-references: UNIPROT:Q81592; UNIPARC:UPI00000BFD24; EMBL:D11353; NID:9221625; PID

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1992

C:Superfamily: hepatitis C virus genome polyprotein

Query Match 5.9%; Score 7; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
|||||
DB 58 PDKEVLY 64

RESULT 25

B70647

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) I chain nuoA - Mycobacterium tuberculosis

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: B70647

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70647

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-128 <COL>

A:Cross-references: UNIPROT:P95181; UNIPARC:UPI00001309D1; GB:Z83867; GB:AL123456; NID:9

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: nuoA

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 3

C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 5.9%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
DB 9 VLAALAA 15

RESULT 26

- RESULT 28

C72737

hypothetical protein APE0433 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: C72737

D95864

hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasamid pSym

C:Species: Sinorhizobium meliloti

C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C:Accession: D95864

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna;

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1.683-kb pSymB megaplasamid from the N2-fixing endo

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: D95864

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-159 <KUR>

A:Cross-references: UNIPROT:Q92WY2; UNIPARC:UPI00000CB440; GB:AL591985; PIDN:CAC48580.1

A:Experimental source: strain 1021, megaplasamid pSymB

R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SWB20180

A:Genome: plasmid

Query Match 5.9%; Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
DB 17 VLAALAA 23

RESULT 27

B84260

hypothetical protein Vng1036h [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: B84260

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,

; Leitchauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl

Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L

A:Title: Genome sequence of Halobacterium species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: B84260

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-172 <STO>

A:Cross-references: UNIPROT:Q9HQR7; UNIPARC:UPI0000637C4; GB:AE004437; NID:g10580588;

C:Genetics:

A:Gene: VNG1036H

Query Match 5.9%; Score 7; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
DB 8 VLAALAA 14

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum

A;Accession: C72737

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-174 <KAW>

A;Cross-references: UNIPROT:Q9YF05; UNIPARC:UPI000005DB3C; DDBJ:AP0000059; NID:G5103911;

A;Experimental source: strain K1

C;Genetics:

C;Superfamily: uncharacterized conserved protein MJ1600

Query Match 5.9%; Score 7; DB 2; Length 174;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25

Db 73 GVLAALA 79

RESULT 29

AE3363

transcription regulator, tetr family BMEI0891 [imported] - Brucella melitensis (strain 1

C;Species: Brucella melitensis

C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C;Accession: AE3363

R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,

; Mazur, M.; Goltsman, E.; Selkov, E.; Eizer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A;Reference number: AD3252; PMID:11756688

A;Accession: AE3363

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-188 <KUR>

A;Cross-references: UNIPROT:Q8YHB1; UNIPROT:Q8GOK2; UNIPARC:UPI0000057B65; GB:AE008917;

A;Experimental source: strain 16M

C;Genetics:

A;Gene: BMEI0891

A;Map position: 1

Query Match 5.9%; Score 7; DB 2; Length 188;

Best Local Similarity 100.0%; Pred. No. 33;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 5.9%; Score 7; DB 1; Length 203;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 AAPYIEQ 74

Db 190 AAPYIEQ 196

RESULT 31

G95276

probable transcription regulator [imported] - Sinorhizobium meliloti (strain 1021) magap

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C;Accession: G95276

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A;Reference number: A95462; MUID:21396509; PMID:11481432

A;Accession: G95276

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-204 <KUR>

A;Cross-references: UNIPROT:Q930S6; UNIPARC:UPI000000CAPAE; GB:AE006469; PIDN:AAK64777.1;

A;Experimental source: strain 1021, megaplasmid pSymA

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,

pela, D.; Chalm, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandendol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMA0223

A;Genome: plasmid

Query Match 5.9%; Score 7; DB 2; Length 204;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25

Db 119 GVLAALA 125

RESULT 32

B81067

conserved hypothetical protein NMB1578 [imported] - Neisseria meningitidis (strain MC58

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: B81067

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizzi, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: B81067

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-217 <TET>

A;Cross-references: UNIPROT:Q9JVH9; UNIPARC:UPI000000C473F; GB:AE002508; GB:AE002098; NID

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB1578

Query Match 5.9%; Score 7; DB 2; Length 217;

Best Local Similarity 100.0%; Pred. No. 37;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 14 VLAALAA 20

RESULT 33
G81801
Probable lipoprotein NMA1767 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C/Species: Neisseria meningitidis
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: G81801
R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A/Reference number: A81775; MUID:20222556; PMID:10761919
A/Accession: G81801
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-217 <PAR>
A/Cross-references: UNIPROT:Q9UT10; UNIPARC:UPI000000445D0; GB:AL162757; GB:AL157959; NID
A/Experimental source: serogroup A, strain Z2491
C/Genetics:
A/Gene: NMA1767

```

Query Match          5.9%; Score 7; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20  VLAALAA 26
      |||||
DB      14  VLAALAA 20

RESULT 34
B95207
aquaporin [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 05-Oct-2004
C;Accession: B95207
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid  

on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf  

nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,  

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: B95207
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-222 <KUR>
A;Cross-references: UNIPROT:Q97P66; UNIPARC:UPI000000519A8; GB:AE005672; PIDN:AAK75851.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SPI778
C;Superfamily: aquaporin

```

```

Query Match          5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      18 GGVLAAL 24
        |||||
Db      206 GGVLAAL 212

RESULT 35
B98072
aquaporin Z, water channel protein [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 05-Oct-2004
C;Accession: B98072

```

R; Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Bargett, S.; DeHoff, B.S.; E. R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; Meyer, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A: Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.,
A: Title: Genome of the Bacterium *Streptococcus pneumoniae* Strain R6.
A: Reference number: A597872; MOID:21429245; PMID:11544234
A: Accession: B98072
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-222 <KUR>
A: Cross-references: UNIPROT:Q8DNP7; UNIPARC:UPI00000E36AC; GB:AE007317; PIDN:AAL00407.1;
C: Genetics:
A: Gene: aqpZ
C: Superfamily: aquaporin

```
Query Match          5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	18	GGVLAAL	24
Db	206	GGVLAAL	212

RESULT 36

RESULT 36
 C83663
 hypothetical protein BH0107 [imported] - *Bacillus halodurans* (strain C-125)
 C/Species: *Bacillus halodurans*
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C/Accession: C83663
 R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: C83663
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-228 <STO>
 A/Cross-references: UNIPROT:Q9KGf8; UNIPARC:UPI000012D915; GB:AF001507; GB:BA000004; NID
 A/Experimental source: strain C-125

C;Superfamily: 4-diphosphocytidyl-2-methyl-D-erythritol synthase

```
Query Match          5.9%; Score 7; DB 2; Length 228;
Best Local Similarity 100.0%; Pred.No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 40 IELGGKP 46
Db 25 IELGGKP 31

RESULT 37

```

RESULT 37
D96932
ABC transporter, ATP-binding protein CAC0266 [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 05-Oct-2004
C:Accession: D96932
R:Nolling, J.; Brston, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A:Reference number: A96900; PMID:21359325; PMID:21359325

```

A; Status: preliminary

A;Molecule type: DNA
A;Residues: 1-234 <KUR>
A;Cross-references: UNIPROT:Q97ND2; UNIPARC:UPI00000C9DEC; GB:AE001437; PIDN:AAK78247.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0266

Query Match 5.9%; Score 7; DB 2; Length 234;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Gaps 0;

Qy 84 GKVLGLL 90
 Db 31 GKVLGLL 37

RESULT 38
 E65082
 hypothetical protein b2966 - Escherichia coli (strain K-12)
 C;Species: Escherichia coli
 C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C;Accession: E65082
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
 A;Title: The complete genome sequence of Escherichia coli K-12.
 A;Reference number: A64720; MUID:97426617; PMID:9278503
 A;Accession: E65082
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-235 <BLAT>
 A;Cross-references: UNIPROT:Q46831; UNIPARC:UPI000013BEF8; GB:AE000379; GB:U00096; NID:9
 A;Experimental source: strain K-12, substrain MG1655
 C;Superfamily: conserved hypothetical protein b2966

Query Match 5.9%; Score 7; DB 2; Length 235;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
 Db 12 VLLGGVL 18

RESULT 39
 AG2300
 hypothetical protein alr3958 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AG2300
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-245 <KUB>
 A;Cross-references: UNIPROT:Q8Y077; UNIPARC:UPI00000CE97A; GB:BA000019; PIDN:BA075657.1;
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: alr3958
 C;Superfamily: Synechocystis hypothetical protein sll1461

Query Match 5.9%; Score 7; DB 2; Length 245;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 KPAIVPD 51
 Db 67 KPAIVPD 73

RESULT 40
 JC5671
 methionyl aminopeptidase (EC 3.4.11.18) type II [validated] - Pyrococcus furiosus
 N/Alternate names: methionine aminopeptidase; peptidase M

C;Species: Pyrococcus furiosus
 C;Date: 20-Nov-1997 #sequence_revision 20-Nov-1997 #text_change 09-Jul-2004
 C;Accession: JC5671; PC4494
 R;Tsunasawa, S.; Izu, Y.; Miyagi, M.; Kato, I.
 J. Biochem. 122, 843-850, 1997
 A;Title: Methionine aminopeptidase from the hyperthermophilic archaeon Pyrococcus furiosus
 A;Reference number: JC5671; MUID:98060511; PMID:9399590
 A;Accession: JC5671
 A;Molecule type: DNA
 A;Residues: 1-295 <TSU>
 A;Cross-references: UNIPROT:P56218; UNIPARC:UPI0000112F22
 A;Accession: PC4494
 A;Molecule type: protein
 A;Residues: 1-295 <TS2>
 A;Cross-references: UNIPARC:UPI0000112F22
 C;Function:
 A;Description: (EC 3.4.11.18) [validated, MUID:98060511]; catalyzes the removal of amino
 A;Note: half-life of approximately 4.5 h at 90 degrees C; optimum temperature of around
 C;Superfamily: Escherichia coli methionyl aminopeptidase
 C;Keywords: aminopeptidase; cobalt; metalloprotein; protein biosynthesis

Query Match 5.9%; Score 7; DB 2; Length 295;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47
 Db 43 ELGGKPA 49

RESULT 41
 C71107
 probable methionyl aminopeptidase - Pyrococcus horikoshii
 C;Species: Pyrococcus horikoshii
 C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
 C;Accession: C71107
 R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekino, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
 A;Reference number: A71000; MUID:98344137; PMID:9679194
 A;Accession: C71107
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-295 <KAW>
 A;Cross-references: UNIPROT:O58362; UNIPARC:UPI0000125A07; GB:AP000003; NID:G3236130; PI:
 A;Experimental source: strain OT3
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C;Genetics:
 A;Gene: PH0628

Query Match 5.9%; Score 7; DB 2; Length 295;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47
 Db 43 ELGGKPA 49

RESULT 42
 AD3473
 methyltransferase (EC 2.1.1.-) [imported] - Brucella melitensis (strain 16M)
 C;Species: Brucella melitensis
 C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C;Accession: AD3473
 R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, .; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A;Reference number: AD3252; PMID:11756688
 A;Accession: AD3473
 A;Status: preliminary

A:Molecule type: DNA
A:Residues: 1-303 <KUR>
A:Cross-references: UNIPROT:Q9YEV7; UNIPARC:UPI00000581CE; GB:AE008917; PIDN:AAL52951.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11770
A:Map position: 1
C:Superfamily: methyltransferase, YraL type
C:Keywords: methyltransferase

Query Match 5.9%; Score 7; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
Db 139 VLAALAA 145

RESULT 43
F72606
probable methanol dehydrogenase regulator APE1317 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F72606
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72606
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-306 <KAW>
A:Cross-references: UNIPROT:Q9YCE0; UNIPARC:UPI000005DEDL; DDBJ:AF000061; NID:G5104821;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1317
C:Superfamily: methanol dehydrogenase regulatory protein

Query Match 5.9%; Score 7; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
Db 26 VLAALAA 32

RESULT 44
F70603
hypothetical protein Rv1011 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: F70603
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70603
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-306 <COL>
A:Cross-references: UNIPROT:O05596; UNIPARC:UPI000012D94B; GB:Z94752; GB:AL123456; NID:9
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv1011
C:Superfamily: 4-diphosphocytidyl-2C-methyl-D-erythritol 2-phosphate synthase

Query Match 5.9%; Score 7; DB 2; Length 306;

Best Local Similarity 100.0%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
Db 206 VLAALAA 212

RESULT 45
F82541
dihydroorotate dehydrogenase XF2571 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82541
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: For a complete list of authors see reference number A59328 below
A:Accession: F82541
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-359 <SIM>
A:Cross-references: UNIPROT:Q9PAE7; UNIPARC:UPI00000C2AA6; GB:AE004064; GB:AE003849; NID:R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carrer, I.; as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohchad, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, P.; P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Itohako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF2571
C:Superfamily: dihydroorotate oxidase

Query Match 5.9%; Score 7; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 TQQQAVI 100
|||||
Db 111 TQQQAVI 117

RESULT 46
T00882
hypothetical protein At2g45690 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein F17K2.22
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 16-Feb-2001
C:Accession: T00882; F84893
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, March 1998
A:Description: Arabidopsis thaliana chromosome II BAC F17K2 genomic sequence.
A:Reference number: Z14207
A:Accession: T00882
A>Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-360 <ROU>
A:Cross-references: UNIPARC:UPI000017936E; EMBL:AC003680; NID:g2979540; PID:g2979558
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F04893
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <STO>
A:Cross-references: UNIPARC:UPI000017936B; GB:AB002093; NID:G6598429; PIDN:AAC28524.2; C:Genetics:
A:Gene: F17K2.22; At3g45690
A:Map position: 2
A:Introns: 26/1, 99/2, 297/3, 320/2, 341/1
C:Superfamily: *Arabidopsis thaliana* hypothetical protein F17K2.22

Query Match 5.9%; Score 7; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 QQAVIEP 102
Db 192 QQAVIEP 198
|||||

RESULT 47
T34921
probable methyltransferase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34921
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z21558
A:Accession: T34921
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-362 <OLI>
A:Cross-references: UNIPROT:O54158; UNIPARC:UPI00000DABA4; EMBL:AL021409; PIDN:CAA16186.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC3F7.15

Query Match 5.9%; Score 7; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 60 GVLAALA 66
|||||

RESULT 48
S77203
hypothetical protein slr1821 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S77203
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77203
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-366 <KAN>
A:Cross-references: UNIPROT:P73714; UNIPARC:UPI0000032C9F; EMBL:D90908; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 5.9%; Score 7; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 3 VLAALAA 9
|||||

RESULT 49
B72594
hypothetical protein APE1219 - *Aeropyrum pernix* (strain KI)
C:Species: *Aeropyrum pernix*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B72594
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72594
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-370 <KAW>
A:Cross-references: UNIPROT:Q9YCN9; UNIPARC:UPI000005DE6D; DBJ:AP000061; NID:G5104821;
A:Experimental source: strain KI
C:Genetics:
A:Gene: APE1219

Query Match 5.9%; Score 7; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 342 VLAALAA 348
|||||

RESULT 50
F88486
Protein F20H11.5 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F88486
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog
A:Reference number: A75000; MUID:99069613; PMID:9831916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: F88486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-383 <STO>
A:Cross-references: UNIPROT:O01739; UNIPARC:UPI00000788D5; GB:chr_III; PIDN:AAB53982.1;
A:Note: Similar to D-amino acid oxidase
C:Genetics:
A:Gene: F20H11.5
A:Map position: 3
C:Superfamily: D-amino-acid oxidase

Query Match 5.9%; Score 7; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 22
Db 209 LGGVLA 215
|||||

RESULT 51
F86268
aminoalcoholphosphotransferase [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 05-Oct-2004
C:Accession: F86268
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86268
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-389 <STO>
 A:Cross-references: UNIPROT:082567; UNIPARC:UPI00000AB43E; GB:AE005172; NID:g9802754; P1
 C:Genetics:
 A:Map position: 1
 C:Superfamily: choline/ethanolaminephosphotransferase

Query Match 5.9%; Score 7; DB 2; Length 389;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVLT 21
 DB 264 VLLGGVLT 270

RESULT 52
 T14412
 ethanolaminephosphotransferase homolog - turnip
 N:Alternate names: aminoalcoholphosphotransferase
 C:Species: Brassica rapa (turnip)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 05-Oct-2004
 C:Accession: T14412
 R:Min, K.M.; Cho, S.H.
 submitted to the EMBL Data Library, April 1997
 A:Reference number: Z18076
 A:Accession: T14412
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-389 <MIN>
 A:Cross-references: UNIPROT:004178; UNIPARC:UPI000009FD47; EMBL:U96713; NID:g2072355; P1
 A:Experimental source: strain L, var. pekinensis Makino
 C:Genetics:
 A:Gene: AAPT1
 C:Superfamily: choline/ethanolaminephosphotransferase

Query Match 5.9%; Score 7; DB 2; Length 389;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVLT 21
 DB 264 VLLGGVLT 270

RESULT 53
 G83413
 Probable MFS transporter PA1848 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: G83413
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.N.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: G83413
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-389 <STO>

A:Cross-references: UNIPROT:Q912P7; UNIPARC:UPI00000C54CC; GB:AE004611; GB:AE004091; N1
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1848

Query Match 5.9%; Score 7; DB 2; Length 389;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
 DB 140 GGVLAAL 146

RESULT 54
 A72698
 hypothetical protein APE1001 - Aeropyrum pernix (strain K1)
 C:Species: Aeropyrum pernix
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: A72698
 R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaiawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix.
 A:Reference number: A72450; MUID:99310339; PMID:10382966
 A:Accession: A72698
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-390 <KAW>
 A:Cross-references: UNIPROT:Q9YDB2; UNIPARC:UPI00005DDDB8; DDBJ:AF000060; NID:g5104188;
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE1001

Query Match 5.9%; Score 7; DB 2; Length 390;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 DB 266 VLAALAA 272

RESULT 55
 B75290
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: B75290
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: B75290
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-396 <WHI>
 A:Cross-references: UNIPROT:Q9RS23; UNIPARC:UPI00000C1ACA; GB:AE002062; GB:AE000513; N1
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2304
 A:Map position: 1
 C:Superfamily: tetracycline resistance protein

Query Match 5.9%; Score 7; DB 2; Length 396;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
 DB 160 LLGGVLA 166

RESULT 56
AI3554
xylase transport system permease protein xylH [imported] - Brucella melitensis (strain 1
C/Species: Brucella melitensis
C/Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C/Accession: AI3554
R/DelVecchio, V.G.; Kapur, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A/Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A/Reference number: AD3252; PMID:11756688
A/Accession: AI3554
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-396 <KUR>
A/Cross-references: UNIPROT:Q8YD15; UNIPARC:UPI0000058459; GB:AE008918; PIDN:AAL53604.1;
A/Experimental source: strain 16M
C/Genetics:
A/Gene: BMEI10362
A/Map position: II
C/Superfamily: 1-arabinose transport system permease araH

Query Match 5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
|||||
Db 301 GVLAALA 307

RESULT 57
C84176
oxalate/formate antiporter [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: C84176
R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: C84176
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-410 <STO>
A/Cross-references: UNIPROT:Q9HSM8; UNIPARC:UPI0000063575; GB:AE004437; NID:g10579805; E
C/Genetics:
A/Gene: oxt
C/Superfamily: hypothetical protein c0103

Query Match 5.9%; Score 7; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 22
|||||
Db 389 LGGVLA 395

RESULT 58
D82487
multidrug resistance protein D VCA0214 [imported] - Vibrio cholerae (strain N16961 serog
C/Species: Vibrio cholerae
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: D82487
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A/Reference number: A82035; MUID:20406833; PMID:10952301
A/Accession: D82487
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-414 <HEI>
A/Cross-references: UNIPROT:Q9KMW4; UNIPARC:UPI00000C347E; GB:AE004361; GB:AE003853; NID
A/Experimental source: serogroup O1; strain N16961; biotype El Tor
C/Genetics:
A/Gene: VCA0214
A/Map position: 2

Query Match 5.9%; Score 7; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
|||||
Db 247 VLLGGVL 253

RESULT 59
T11387
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - banded wood snail mitochondrion
C/Species: mitochondrion Cepaea nemoralis (banded wood snail)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11387
R/Yamazaki, N.; Ueshima, R.; Terrett, J.A.; Yokobori, S.; Kaifu, M.; Segawa, R.; Kobayas
Genetics 145, 749-758, 1997
A/Title: Evolution of pulmonate gastropod mitochondrial genomes: Comparisons of complete
A/Reference number: Z17266; MUID:97207848; PMID:9055084
A/Accession: T11387
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-417 <YAM>
A/Cross-references: UNIPROT:Q34184; UNIPARC:UPI000008C4F6; EMBL:U23045; NID:g2160231; PI
A/Experimental source: hepatopancreas
C/Genetics:
A/Genome: mitochondrion
C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: mitochondrion; NAD; oxidoreductase

Query Match 5.9%; Score 7; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
|||||
Db 235 GGVLAAL 241

RESULT 60
B70753
probable efflux protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: B70753
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: B70753
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-419 <COL>
A/Cross-references: UNIPROT:Q11060; UNIPARC:UPI000013A501; GB:Z77137; GB:AL123456; NID:G
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: Rv1258c

RESULT 65

C81871

probable histidine-tRNA ligase (EC 6.1.1.21) NMA1065 [imported] - Neisseria meningitidis

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C;Accession: G81871

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jegelse, K.; Leather, S.; Moulé, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: G81871

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-431 <PAR>

A;Cross-references: UNIPROT:Q9JUZ9; UNIPARC:UPI00000C4AEC; GB:ALI62755; GB:ALI57959; NID

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: hls9, NMA1065

C;Superfamily: histidyl-tRNA synthetase; histidine-tRNA ligase homology

C;Keywords: ligase

Query Match 5.9%; Score 7; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47

Db 302 ELGGKPA 308

|||||

RESULT 66

A81150

histidyl-tRNA synthetase NMB0854 [imported] - Neisseria meningitidis (strain MC58 serogr

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004

C;Accession: A81150

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: A81150

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-431 <TET>

A;Cross-references: UNIPROT:Q9JZX9; UNIPARC:UPI00000C45AA; GB:AE002439; GB:AE002098; NID

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB0854

C;Superfamily: histidyl-tRNA synthetase; histidine-tRNA ligase homology

Query Match 5.9%; Score 7; DB 2; Length 431;

Best Local Similarity 100.0%; Pred. No. 64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47

Db 302 ELGGKPA 308

|||||

RESULT 67

E70590

3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) - Mycobacterium tuberculosis

N;Alternate names: arC protein

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C;Accession: E70590; A37807

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: E70590

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-450 <COL>

A;Cross-references: UNIPROT:P22487; UNIPARC:UPI0000125F45; GB:Z95121; GB:ALI123456; NID:9

A;Experimental source: strain H37Rv

R;Garbe, T.; Jones, C.; Charles, I.; Dougan, G.; Young, D.

J. Bacteriol. 172, 6774-6782, 1990

A;Title: Cloning and characterization of the arC gene from Mycobacterium tuberculosis.

A;Reference number: A37807; MUID:91072223; PMID:2123856

A;Accession: A37807

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-450 <GAR>

A;Cross-references: UNIPARC:UPI0000125F45; GB:M62708; NID:G149927; PIDN:AAA25356.1; PID:

C;Genetics:

A;Gene: arC

C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox

C;Keywords: transferase

F;15-417/Domain: 3-phosphoshikimate 1-carboxyvinyltransferase homology <PSK>

Query Match 5.9%; Score 7; DB 2; Length 450;

Best Local Similarity 100.0%; Pred. No. 67;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 31 VLAALAA 37

|||||

RESULT 68

C82066

magnesium transporter VC2534 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C;Species: Vibrio cholerae

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: C82066

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.I.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: C82066

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-453 <HEI>

A;Cross-references: UNIPROT:Q9KPA5; UNIPARC:UPI00000C3315; GB:AE004322; GB:AE003852; NID

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC2534

C;Superfamily: magnesium transport protein mgE

Query Match 5.9%; Score 7; DB 2; Length 453;

Best Local Similarity 100.0%; Pred. No. 67;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 296 VLAALAA 302

|||||

RESULT 69

C82825

UDP-N-acetylmuramate-L-alanine ligase XF0276 [imported] - Xylella fastidiosa (strain 9a5

C;Species: Xylella fastidiosa

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: C82825

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: C82825
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-471 <SIM>
A:Cross-references: UNIPROT:Q9PGM2; UNIPARC:UPI00000C2362; GB:AE003881; GB:AE003849; NID
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GeneBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tauhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0276
C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 5.9%; Score 7; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
DB 306 VLAALAA 312

RESULT 70
S71862
protein disulfide-isomerase (EC 5.3.4.1) precursor - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 05-May-2000
C:Accession: S71862
R:Veijola, J.; Annunen, P.; Koivunen, P.; Page, A.P.; Pihlajaniemi, T.; Kivirikko, K.I. Biochem. J. 317, 721-729, 1996
A:Title: Baculovirus expression of two protein disulphide isomerase isoforms from Caenor ta subunit.
A:Reference number: S71862; MUID:96332462; PMID:8760355
A:Accession: S71862
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-491 <VEI>
A:Cross-references: UNIPARC:UPI00001761A5
A:Experimental source: lumen of endoplasmic reticulum
C:Genetics:
A:Gene: pdi
A:Map position: III
C:Complex: forms an alpha-beta dimer with procollagen-proline dioxygenase (EC 1.14.11.2)
C:Function: <F1>
A:Description: catalyzes rearrangement of both intrachain and interchain disulfide bonds
C:Function: <F2>
A:Description: acts as the beta-subunit of procollagen-proline dioxygenase (EC 1.14.11.2)
C:Function: <F3>
A:Description: acts as a chaperone-like protein
C:Superfamily: protein disulfide-isomerase; thioredoxin homology
C:Keywords: duplication; endoplasmic reticulum; intramolecular oxidoreductase; isomerase
F:1-14/Domain: signal sequence (fragment) #status predicted <SIG>
F:15-491/Product: protein disulfide-isomerase #status predicted <MAT>
F:29-113/Domain: thioredoxin homology <TX1>
F:369-453/Domain: thioredoxin homology <TX2>
F:488-491/Region: endoplasmic reticulum retention signal
F:50-53/Disulfide bonds: redox-active #status predicted

F:391-394/Disulfide bonds: redox-active #status predicted

Query Match 5.9%; Score 7; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QFKGKVL 87
|||||
DB 275 QFKGKVL 281

RESULT 71
T34092
hypothetical protein C07A12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34092
R:Geisel, C.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid C07A12.
A:Reference number: Z21475
A:Accession: T34092
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-493 <GEI>
A:Cross-references: UNIPROT:Q17770; UNIPARC:UPI0000060EC2; EMBL:U41542; PIDN:AAC69237.1.
A:Experimental source: strain Bristol N2; clone C07A12
C:Genetics:
A:Gene: CESP:C07A12.4
A:Map position: X
A:Introns: 30/3; 158/3; 427/3
C:Superfamily: protein disulfide-isomerase; thioredoxin homology

Query Match 5.9%; Score 7; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QFKGKVL 87
|||||
DB 277 QFKGKVL 283

RESULT 72
AF2325
NADH dehydrogenase chain 4 ndhD [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2325
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2325
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-500 <NUT>
A:Cross-references: UNIPROT:Q8YPN6; UNIPARC:UPI00000CEA34; GB:BA000019; PIDN:BA075856.1.
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: ndhD
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

Query Match 5.9%; Score 7; DB 2; Length 500;
Best Local Similarity 100.0%; Pred. No. 73;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
|||||
DB 243 LLGGVLA 249

RESULT 73

G72464
hypothetical protein APE2362 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 05-Oct-2004
C:Accession: G72464
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; KDNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: G72464
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-510 <KAW>
A:Cross-references: UNIPROT:Q9Y9CS; UNIPARC:UPI000005E2FB; DDBJ:AP0000064; NID:gs105945;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2362
C:Superfamily: uncharacterized conserved protein

Query Match 5.9%; Score 7; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 27

Db 280 LAALAA 286

RESULT 74

A87482
conserved hypothetical protein CC1878 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87482
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87482
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-510 <STO>
A:Cross-references: UNIPROT:Q9A748; UNIPARC:UPI00000C7533; GB:AE005673; NID:gl3423323; E
C:Genetics:
A:Gene: CC1878

Query Match 5.9%; Score 7; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25

Db 179 GVLAALA 185

RESULT 75

F96829
hypothetical protein F19K16.18 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: F96829
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Mazziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96829
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-515 <STO>
A:Cross-references: UNIPROT:Q9CA89; UNIPARC:UPI000009F627; GB:AE005173; NID:g6453861; P
C:Genetics:
A:Gene: F19K16.18
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein T7H20.60

Query Match 5.9%; Score 7; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 75 VLAALAA 81

RESULT 76

D82063
Probable thiamin ABC transporter, permease protein VC2538 [imported] - Vibrio cholerae
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82063
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82063

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-530 <HBI>

A:Cross-references: UNIPROT:Q9KP41; UNIPARC:UPI00000C3319; GB:AE004323; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2538

A:Map position: 1

Query Match 5.9%; Score 7; DB 2; Length 530;

Best Local Similarity 100.0%; Pred. No. 76;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 19 VLAALAA 25

RESULT 77

Q9ECD7

tnsE protein - Escherichia coli transposon Tn7

C:Species: Escherichia coli

C>Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004

C:Accession: A25543; S12641; S06770

R:Smith, G.M.; Jones, P.

Nucleic Acids Res. 14, 7915-7927, 1986

A:Title: Tn7 transposition: a multigenic process. Identification of a regulatory gene pro

A:Reference number: A93644; MUID:87040763; PMID:3022239

A:Accession: A25543

A:Molecule type: DNA

A:Residues: 1-538 <SMI>

A:Cross-references: UNIPROT:P05845; UNIPARC:UPI0000000F5C; GB:X04534; NID:g43752; PIDM:C

R:Flores, C.; Gadi, M.I.; Lichtenstein, C.

Nucleic Acids Res. 18, 901-911, 1990

A:Title: DNA sequence analysis of five genes; tnsA, B, C, D and E, required for Tn7 tran

A:Reference number: S12637; MUID:90192166; PMID:2156235

A:Accession: S12641

A:Molecule type: DNA

A:Residues: 1-538 <P>LO-
A:Cross-references: UNIPARC:UPI000000095C; EMBL:XL17693; NID:G43755; PIDN:CAA35687.1; PID
A>Note: The authors translated the initiation codon GTG for residue 1 as Val
C:Genetics:
A:Gene: tnsE
A:Start codon: GTG
C:Function:
A:Description: required for the transposition of transposon Tn7
C:Superfamily: tnsE protein
C:Keywords: DNA binding; transposition

Query Match 5.9%; Score 7; DB 1; Length 538;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 23
|||||
DB 363 LGGVLA 369

RESULT 78
E84213
hypothetical protein Vng0553c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84213
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A>Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: E84213
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-559 <STO>
A:Cross-references: UNIPROT:Q9HRT5; UNIPARC:UPI0000063675; GB:AE004437; NID:gl0580151; F
C:Genetics:
A:Gene: VNG0553C

Query Match 5.9%; Score 7; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
DB 111 VLAALAA 117

RESULT 79
G91231
probable frv operon regulatory protein [imported] - Escherichia coli (strain O157:H7, su
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: G91231
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: G91231
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <HAY>
A:Cross-references: UNIPROT:Q8X8A7; UNIPARC:UPI000000096C; GB:BA000007; PIDN:BA838246.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC84823

Query Match 5.9%; Score 7; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQR 92
|||||
DB 252 VLGLLQR 258

RESULT 80
F86078

probable frv operon regulatory protein frvr [imported] - Escherichia coli (strain O157:H
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F86078
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F86078
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-582 <STO>
A:Cross-references: UNIPROT:Q8X8A7; UNIPARC:UPI000000096C; GB:AE005174; NID:gl2518792; F
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: frvr

Query Match 5.9%; Score 7; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQR 92
|||||
DB 252 VLGLLQR 258

RESULT 81
H83941

ABC transporter (ATP-binding protein) BH2336 [imported] - Bacillus halodurans (strain C-
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: H83941
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: H83941
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-585 <STO>
A:Cross-references: UNIPROT:Q9KAF1; UNIPARC:UPI00000003EA3; GB:AP001515; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Superfamily: Escherichia coli ABC transporter mdla; ATP-binding cassette homology

Query Match 5.9%; Score 7; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
|||||
DB 61 GVLAALA 67

RESULT 82
H97105

carbon starvation protein [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: H97105
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: H97105
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-592 <KUR>
A;Cross-references: UNIPROT:Q97IH2; UNIPARC:UPI00000CA2A1; GB:AE001437; PTDN:AAK79635.1;
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1669

Query Match 5.9%; Score 7; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 22
| | | | |
Db 90 LGGVLA 96
| | | | |

RESULT 83
H75272
probable nucleic acid-binding protein, HRDC family - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75272
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75272
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-603 <WHI>
A;Cross-references: UNIPROT:Q9RRP4; UNIPARC:UPI00000D3PBA; GB:AE002074; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2444
A;Map position: 1

Query Match 5.9%; Score 7; DB 2; Length 603;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
| | | | |
Db 227 LGGVLA 233
| | | | |

RESULT 84
S74455
ABC-type transport protein slr1488 - Synecocystis sp. (strain PCC 6803)
A;Alternate names: protein slr1488
C;Species: Synecocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 05-Oct-2004
C;Accession: S74455
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74455
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-605 <KAN>
A;Cross-references: UNIPROT:P72607; UNIPARC:UPI00000D31BE; EMBL:D90899; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Keywords: ATP; nucleotide binding; P-loop; transport protein.
F;379-573/Domain: ATP-binding cassette homology <ABC>

F;396-403/Region: nucleotide-binding motif A (P-loop)
Query Match 5.9%; Score 7; DB 2; Length 605;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGLL 90
| | | | |
Db 389 GKVLGLL 395
| | | | |

RESULT 85
F71019
hypothetical protein PH1451 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C;Accession: F71019
R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: F71019
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-617 <KAW>
A;Cross-references: UNIPROT:O59120; UNIPARC:UPI00000631A9; GB:AP0000006; NID:G3236133; P
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH1451

Query Match 5.9%; Score 7; DB 2; Length 617;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
| | | | |
Db 292 GGVLAAL 298
| | | | |

RESULT 86
A75390
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 DR1494 [similarity] - Deinococcus
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75390
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: A75390
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-645 <WHI>
A;Cross-references: UNIPROT:Q9RU98; UNIPARC:UPI00000D3B67; GB:AE001993; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1494
A;Map position: 1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C;Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 5.9%; Score 7; DB 2; Length 645;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 25
| | | | |
Db 475 GVLAAAL 481
| | | | |

RESULT 87
AH0411
ferrichrome transport system permease protein FhuB fhuB [imported] - Yersinia pestis (strain 1021) #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Species: Yersinia pestis
C:Date: 02-Nov-2001
C:Accession: AH0411
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-665 <KUR>
A:Cross-references: UNIPROT:Q8ZBL8; UNIPARC:UPI00000CDA12; GB:AL590842; PIDN:CAC92620.1;
C:Genetics:
A:Gene: fhuB
C:Superfamily: vitamin B12 transport protein btuC

Query Match 5.9%; Score 7; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
|||||
Db 399 VLAALAA 405

RESULT 88
E81410
transcription regulatory protein hypF Cj0622 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 14-Apr-2003
C:Accession: E81410
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Bagham, D.; Chillingworth, T.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrrell Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypA:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: E81410
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-729 <PAR>
A:Cross-references: UNIPARC:UPI00000C21B1; GB:AL139075; GB:AL111168; NID:G6967817; PIDN:G6967817
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: hypF; Cj0622
C:Superfamily: carbamoyl phosphate-converting enzyme ([NiFe]-hydrogenase maturation factor)

Query Match 5.9%; Score 7; DB 2; Length 729;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
|||||
Db 274 LLGGVLA 280

RESULT 89
A95421
probable oxidoreductase SMA2353 [imported] - Sinorhizobium meliloti (strain 1021) magap1
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: A95421
R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe, S.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C. Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: A95421

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-733 <KUR>
A:Cross-references: UNIPROT:Q92KH9; UNIPARC:UPI00000CB384; GB:AE006469; PIDN:AAK65931.1;
R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Pederspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA2353
A:Genome: plasmid
C:Superfamily: carbon monoxide dehydrogenase molybdo-protein

Query Match 5.9%; Score 7; DB 2; Length 733;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
|||||
Db 253 VLAALAA 259

RESULT 90
F71369
conserved hypothetical protein TP0079 - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C:Accession: F71369
R:Praser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin, J.; Khaldun, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo, they, L.; Kladman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: F71369
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-740 <COL>
A:Cross-references: UNIPROT:O83118; UNIPARC:UPI00000C0A43; GB:AE001192; GB:AE000520; NID:O83118
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0079

Query Match 5.9%; Score 7; DB 2; Length 740;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAYC 28
|||||
Db 242 AALAAAYC 248

RESULT 91
D70581
probable cation-transporting ATPase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 31-Dec-2004
C:Accession: D70581
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70581
A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 1-797 <COL>
A:Cross-references: UNIPROT:O08365; UNIPARC:UPI0000128643; GB:Z95210; GB:AL123456; NID:9
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: ctpG
C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom
F:436-601/Domain: ATPase nucleotide-binding domain homology <ATM>
Query Match 5.9%; Score 7; DB 2; Length 797;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLALAA 26
| | | | |
Db 331 VLALAA 337
| | | | |
RESULT 92
I57487
Na+/H+-exchanging protein, amiloride-sensitive - human
N:Alternate names: Na+/H+ antiporter; NHE-1
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
R:Accession: I57487; A31311
R:Flyeig, L.; Dyck, J.R.; Wang, H.; Fong, C.; Haworth, R.S.
Mol. Cell. Biochem. 125, 137-143, 1993
A:Title: Cloning and analysis of the human myocardial Na+/H+ exchanger.
A:Reference number: I57487; MUID:94111706; PMID:8283968
A:Accession: I57487
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-815 <FLI>
A:Cross-references: UNIPROT:P19634; UNIPARC:UPI000012FD1B; GB:S68616; NID:G544775; PIDN:
R:Sardet, C.; Franchi, A.; Pouyssegur, J.
Cell 56, 271-280, 1989
A:Title: Molecular cloning, primary structure, and expression of the human growth factor
A:Reference number: A31311; MUID:89106219; PMID:2536298
A:Accession: A31311
A:Molecule type: mRNA
A:Residues: 1-814, 'SNARASQRLSPHRLPHQSGWGLPLPFLTRIGAPPPTAQWLPQPPQHSPPAASREASSPPELR
A:Cross-references: UNIPARC:UPI000017C2B5
C:Genetics:
A:Gene: GDB:SLC9A1; APNH; NHE1
A:Cross-references: GDB:119683; OMIM:107310
A:Map position: 1p36.1-1p35
C:Keywords: glycoprotein; transmembrane protein
F:75,370,410/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 5.9%; Score 7; DB 2; Length 815;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 83 KGKVLGL 89
| | | | |
Db 738 KGKVLGL 744
| | | | |
RESULT 93
S16328
Na+/H+-exchanging protein - rabbit
N:Alternate names: Na+/H+ antiporter; Na+/H+ exchanger; pH regulatory protein
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
R:Accession: S16328; S13926; S30602
R:Jesse, C.W.; Ma, A.I.; Yang, V.W.; Watson, A.J.M.; Levine, S.; Montrose, M.H.; Potter, J
EMBO J. 10, 1957-1967, 1991
A:Title: Molecular cloning and expression of a cDNA encoding the rabbit ileal villus cel
A:Reference number: S16328; MUID:91293066; PMID:1712287
A:Accession: S16328
A:Molecule type: mRNA
A:Residues: 1-816 <TSB>
A:Cross-references: UNIPROT:P23791; UNIPARC:UPI000012FD1D; EMBL:X59935; NID:G1642; PIDN:

R:Fliegel, L.; Sardet, C.; Pouyssegur, J.; Barr, A.
FEBS Lett. 279, 25-29, 1991
A:Title: Identification of the protein and cDNA of the cardiac Na(+)/H(+) exchanger.
A:Reference number: S13926; MUID:91138752; PMID:1704856
A:Accession: S13926
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 472-816 <FLI>
A:Cross-references: UNIPARC:UPI000016CSA5; EMBL:X56536; NID:G1666; PIDN:CAA39881.1; PID:
R:Hildebrandt, F.; Pizzonia, J.H.; Reilly, R.F.; Reboucas, N.A.; Sardet, C.; Pouyssegur,
Biochim. Biophys. Acta 1129, 105-108, 1991
A:Title: Cloning, sequence, and tissue distribution of a rabbit renal Na(+)/H(+) exchange
A:Reference number: S30602; MUID:92096447; PMID:1661611
A:Accession: S30602
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-241, 'A', 243-568, 'E', 570-816 <HIL>
A:Cross-references: UNIPARC:UPI000016CS9A; EMBL:X61504; NID:G1653; PIDN:CAA43721.1; PID:
A:Note: the authors translated the codon CTG for residue 646 as Val and GCG for residue
C:Keywords: glycoprotein; transmembrane protein
F:16-35/Domain: transmembrane #status predicted <TM1>
F:108-127/Domain: transmembrane #status predicted <TM2>
F:130-149/Domain: transmembrane #status predicted <TM3>
F:155-174/Domain: transmembrane #status predicted <TM4>
F:192-211/Domain: transmembrane #status predicted <TM5>
F:295-315/Domain: transmembrane #status predicted <TM6>
F:339-357/Domain: transmembrane #status predicted <TM7>
F:387-406/Domain: transmembrane #status predicted <TM8>
F:411-430/Domain: transmembrane #status predicted <TM9>
F:481-500/Domain: transmembrane #status predicted <TM10>
F:75,370/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 5.9%; Score 7; DB 2; Length 816;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 83 KGKVLGL 89
| | | | |
Db 738 KGKVLGL 744
| | | | |
RESULT 94
A48858
Na+/H+-exchanging protein - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
R:Accession: A48858; I46613
R:Reilly, R.F.; Hildebrandt, F.; Biemesderfer, D.; Sardet, C.; Pouyssegur, J.; Aronson,
Am. J. Physiol. 261, F1088-F1094, 1991
A:Title: cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in LLC-PK1 renal ep
A:Reference number: A48858; MUID:92087905; PMID:1661081
A:Accession: A48858
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-818 <REI>
A:Cross-references: UNIPROT:P48762; UNIPARC:UPI000016C62E; GB:S71135; NID:G240706; PIDN:
A:Experimental source: LLC-PK1 kidney cell line
A:Note: sequence extracted from NCBI backbone (NCBIN:711135, NCBIP:711136)
R:Reilly, R.F.
Am. J. Physiol. 261, 1088-1094, 1991
A:Title: cDNA cloning and immunolocalization of a Na+-H+ exchanger in LLC-PK1 renal epit
A:Reference number: I46613
A:Accession: I46613
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-682, 'H', 684-818 <RE2>
A:Cross-references: UNIPARC:UPI000012FD1C; GB:M89631; NID:G164595; PIDN:AAA31092.1; PID:
Query Match 5.9%; Score 7; DB 2; Length 818;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 83 KGKVLGL 89

```
Db      738 KGKVLGL 744
|||||
RESULT 95
A40204
Na+/H+-exchanging protein 1 - rat
N:Alternate names: Na+/H+ antiporter
C:Species: Rattus norvegicus (Norway rat)
C>Date: 28-Aug-1992 #sequence_revision 28-Aug-1992 #text_change 28-Feb-1997
C:Accession: A40204
R:Orlowski, J.; Kandasamy, R.A.; Shull, G.B.
J. Biol. Chem. 267, 9331-9339, 1992
A>Title: Molecular cloning of putative members of the Na/H exchanger gene family. cDNA c
rally related proteins.
A:Reference number: A40204; MUID:92250539; PMID:1577762
A:Accession: A40204
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-820 <ORF>
A:Cross-references: UNIPARC:UPI000017C960; GB:M85299
C:Keywords: transmembrane protein

Query Match      5.9%; Score 7; DB 2; Length 820;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      83 KGKVLGL 89
|||||
Db      742 KGKVLGL 748

RESULT 96
F90899
hypothetical protein ECs2166 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: F90899
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ohnishi, M.; Kurokawa, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F90899
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-870 <HAY>
A:Cross-references: UNIPROT:Q8X316; UNIPARC:UPI00000D2978; GB:BA000007; PIDN:BA035589.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs2166
C:Superfamily: phage lambda minor tail protein H

Query Match      5.9%; Score 7; DB 2; Length 870;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 VLAALAA 26
|||||
Db      166 VLAALAA 172

RESULT 97
E90997
hypothetical protein ECs2949 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: E90997
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A:Reference number: A99629; MUID:21156231; PMID:11258796
```

```
A:Accession: E90997
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-881 <HAY>
A:Cross-references: UNIPROT:Q9EYEL; UNIPARC:UPI00000D2952; GB:BA000007; PIDN:BA036372.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs2949
C:Superfamily: phage lambda minor tail protein H

Query Match      5.9%; Score 7; DB 2; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 VLAALAA 26
|||||
Db      177 VLAALAA 183

RESULT 98
F83530
hypothetical protein PA0920 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83530
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bli
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83530
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-881 <STO>
A:Cross-references: UNIPROT:Q9I537; UNIPARC:UPI00000C51BB; GB:AE004526; GB:AE004091; NIT
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA0920

Query Match      5.9%; Score 7; DB 2; Length 881;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 VLAALAA 26
|||||
Db      166 VLAALAA 172

RESULT 99
D85817
hypothetical protein Z3084 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85817
R:Perna, N.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85817
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-884 <STO>
A:Cross-references: UNIPROT:Q8X327; UNIPARC:UPI00000D0E78; GB:AE005174; NID:912516101; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z3084
C:Superfamily: phage lambda minor tail protein H

Query Match      5.9%; Score 7; DB 2; Length 884;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 20 VLAALAA 26
    |||||||
Db 180 VLAALAA 186

RESULT 100
Hypothetical protein (whiB 3' region) - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S18942; S20913
R:Davis, N.K.; Chater, K.F.
submitted to the EMBL Data Library, September 1991
A:Reference number: S18941
A:Accession: S18942
A:Molecule type: DNA
A:Residues: 1-914 <DAV1>
A:Cross-references: UNIPROT:Q53964; UNIPARC:UPI00000839F1; EMBL:X62287; NID:g49001; PIDN:
T51947
R:Davis, N.K.; Chater, K.F.
Mol. Gen. Genet. 232, 351-358, 1992
A:Title: The Streptomyces coelicolor whiB gene encodes a small transcription factor-like
A:Reference number: S20912; MUID:92269753; PMID:1316997
A:Accession: S20913
A:Molecule type: DNA
A:Residues: 1-11 <DAV2>
A:Cross-references: UNIPARC:UPI0000179E96; EMBL:X62287

Query Match 5.9%; Score 7; DB 2; Length 914;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
    |||||||
Db 719 VLAALAA 725

RESULT 101
FXLNPD
H+-exporting ATPase (EC 3.6.3.6), plasma membrane - Leishmania donovani
N:Alternate names: proton-transporting ATPase
C:Species: Leishmania donovani
C>Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 31-Dec-2004
C:Accession: A27124
R:Meade, J.C.; Shaw, J.; Lenaster, S.; Gallagher, G.; Stringer, J.R.
Mol. Cell. Biol. 7, 3937-3946, 1987
A:Title: Structure and expression of a tandem gene pair in Leishmania donovani that encod
A:Reference number: A27124; MUID:88122116; PMID:2828921
A:Accession: A27124
A:Molecule type: DNA
A:Residues: 1-374 <MEA>
A:Cross-references: UNIPARC:UPI0000172EA6; GB:M17889; NID:g159291; PIDN:AAA29227.1; PID:
A:Note: the authors translated the codon AGA for residue 352 as Lys
C:Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom
C:Keywords: ATP; hydrogen ion transport; hydrolase; membrane protein; phosphoprotein
F:495-670/Domain: ATPase nucleotide-binding domain homology <ATN>
F:351/Active site: Asp (aspartylphosphate intermediate) #status predicted

Query Match 5.9%; Score 7; DB 1; Length 974;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
    |||||||
Db 381 VLAALAA 387

RESULT 102
S61236
major capsid protein - bovine herpesvirus 1
C:Species: bovine herpesvirus 1
C>Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004
C:Accession: S61236
```

```
R;Vlcek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Sch
submitted to the EMBL Data Library, January 1995
A:Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus
A:Reference number: S61233
A:Accession: S61236
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1385 <VLC>
A:Cross-references: UNIPROT:Q65565; UNIPARC:UPI000000EBD00; EMBL:Z48053; NID:g971311; PIDN:
C:Superfamily: varicella-zoster virus major capsid protein

Query Match 5.9%; Score 7; DB 2; Length 1385;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 22
    |||||||
Db 288 LGGVLA 294

RESULT 103
T51947
probable transcription factor HUA2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C:Accession: T51947
R:Chen, X.; Meyerowitz, E.M.
Mol. Cell 3, 349-360, 1999
A:Title: HUA1 and HUA2 are two members of the floral homeotic AGAMOUS pathway.
A:Reference number: Z25882; MUID:99214850; PMID:10198637
A:Accession: T51947
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1392 <CHE>
A:Cross-references: UNIPROT:Q9XER9; UNIPARC:UPI000000A0734; EMBL:AF116556; PIDN:AAD31171.
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: HUA2
A:Map position: 5

Query Match 5.9%; Score 7; DB 2; Length 1392;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIVP 50
    |||||||
Db 453 GKPAIVP 459

RESULT 104
A45344
immediate-early protein - suid herpesvirus 1 (strain Kaplan)
C:Species: suid herpesvirus 1
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: A45344
R;Vlcek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzler, M.
Virology 179, 365-377, 1990
A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented op
A:Reference number: A45344; MUID:91021039; PMID:2171211
A:Accession: A45344
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-1446 <VLC>
A:Cross-references: UNIPROT:P33479; UNIPARC:UPI000012D219; GB:M34651; NID:g334070; PIDN:
C:Superfamily: herpesvirus immediate-early protein IE175
C:Keywords: DNA binding; early protein; transcription regulation

Query Match 5.9%; Score 7; DB 1; Length 1446;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAVC 28
    |||||||
```

Db 991 AALAYC 997

RESULT 105

EDBEIF

immediate-early protein IE180 - suid herpesvirus 1 (strain Indiana-Funkhauser)

C:Species: suid herpesvirus 1

C>Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 09-Jul-2004

C:Accession: S04713

R:Cheung, A.K.

Nucleic Acids Res. 17, 4637-4646, 1989

A:Title: DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus

A:Reference number: S04713; MUID:89315207; PMID:2546124

A:Accession: S04713

A:Molecule type: DNA

A:Residues: 1-1460 <CHE>

A:Cross-references: UNIPROT:P11675; UNIPARC:UPI000017497B

C:Superfamily: herpesvirus immediate-early protein IE175

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 5.9%; Score 7; DB 1; Length 1460;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAYC 28

Db 1004 AALAYC 1010

RESULT 106

T00076

hypothetical protein KIAA0462 - human (fragment)

N:Alternate names: hypothetical protein DKFZp564B122.1

C:Species: Homo sapiens (man)

C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004

C:Accession: T00076; T08689

R:Seiki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.;

DNA Res. 4, 345-349, 1997

A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human

A:Reference number: Z14085; MUID:98116662; PMID:9455484

A:Accession: T00076

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-2276 <SEK>

A:Cross-references: UNIPROT:O75050; UNIPARC:UPI000007380B; EMBL:AB007931; NID:G3413885;

A:Experimental source: brain

R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16470

A:Accession: T08689

A:Molecule type: mRNA

A:Residues: 1779-2276 <DUE>

A:Cross-references: UNIPARC:UPI000016AC4A; EMBL:AL049972

A:Experimental source: fetal brain; clone DKFZp564B122

C:Genetics:

A:Map position: 1

A:Note: KIAA0462

A:Note: DKFZp564B122.1

Query Match 5.9%; Score 7; DB 2; Length 2276;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26

Db 425 VLAALAA 431

RESULT 107

JQ1303

genome polyprotein - hepatitis C virus (isolate HC-J6)

N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004

C:Accession: JQ1303

R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.

J. Gen. Virol. 72, 2697-2704, 1991

A:Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human

A:Reference number: JQ1303; MUID:92044440; PMID:1658196

A:Accession: JQ1303

A:Molecule type: genomic RNA

A:Residues: 1-3033 <OKA>

A:Cross-references: UNIPROT:P26660; UNIPARC:UPI0000131E35; GB:D00944; NID:G221650; PIDN

A:Experimental source: isolate HC-J6 from a Japanese individual

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transmembrane

F:116-191/Product: envelope protein M #status predicted <CPC>

F:192-389/Product: major envelope protein E #status predicted <EPM>

F:390-733/Product: nonstructural protein NS1 #status predicted <NS1>

F:734-1010/Product: nonstructural protein NS2 #status predicted <NS2>

F:1011-1619/Product: hepacivirin #status predicted <NS3>

F:1316-1321/Region: nucleotide-binding motif B

F:1320-1323/Region: DEXH motif

F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>

F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>

F:2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>

F:196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,2

Query Match 5.9%; Score 7; DB 1; Length 3033;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56

Db 1700 PDKEVLY 1706

RESULT 108

S09026

carboxylesterase (EC 3.1.1.1) MK1, microsomal - crab-eating macaque (fragment)

C:Species: Macaca fascicularis (crab-eating macaque)

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 09-Jul-2004

C:Accession: S09026

R:Hosokawa, M.; Maki, T.; Satoh, T.

Arch. Biochem. Biophys. 277, 219-227, 1990

A:Title: Characterization of molecular species of liver microsomal carboxylesterases of

A:Reference number: S09021; MUID:90179180; PMID:2310190

A:Accession: S09026

A:Molecule type: protein

A:Residues: 1-18 <HOS>

A:Cross-references: UNIPROT:Q7M2Q1; UNIPARC:UPI000017C057

C:Keywords: carboxylic ester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 45;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGKVLG 88

Db 12 KGKVLG 17

RESULT 109

S09025

carboxylesterase (EC 3.1.1.1), microsomal - rabbit (fragment)

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998

C:Accession: S09025

R:Hosokawa, M.; Maki, T.; Satoh, T.

Arch. Biochem. Biophys. 277, 219-227, 1990

A:Title: Characterization of molecular species of liver microsomal carboxylesterases of

A:Reference number: S09021; MUID:90179180; PMID:2310190

A:Accession: S09025

A:Molecule type: protein

A;Residues: 1-20 <HOS>
A;Cross-references: UNIPARC:UPI0000175900
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
Db 12 KGVVLG 17

RESULT 110

S09023
carboxylesterase (EC 3.1.1.1) RL2, microsomal - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998
C;Accession: S09023

R;Hosokawa, M.; Maki, T.; Satoh, T.

Arch. Biochem. Biophys. 277, 219-227, 1990
A;Title: Characterization of molecular species of liver microsomal carboxylesterases of

A;Reference number: S09021; MUID:90179180; PMID:2310190

A;Accession: S09023

A;Molecule type: protein

A;Residues: 1-20 <HOS>

A;Cross-references: UNIPARC:UPI0000175901

C;Superfamily: cholinesterase; cholinesterase homology

C;Keywords: carboxylic ester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
Db 12 KGVVLG 17

RESULT 111

S09022
carboxylesterase (EC 3.1.1.1) RL1, microsomal - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998
C;Accession: S09022

R;Hosokawa, M.; Maki, T.; Satoh, T.

Arch. Biochem. Biophys. 277, 219-227, 1990
A;Title: Characterization of molecular species of liver microsomal carboxylesterases of

A;Reference number: S09021; MUID:90179180; PMID:2310190

A;Accession: S09022

A;Molecule type: protein

A;Residues: 1-20 <HOS>

A;Cross-references: UNIPARC:UPI00001759F8

C;Superfamily: cholinesterase; cholinesterase homology

C;Keywords: carboxylic ester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
Db 12 KGVVLG 17

RESULT 112

S09021
carboxylesterase (EC 3.1.1.1) RL1, microsomal - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998
C;Accession: S09021

R;Hosokawa, M.; Maki, T.; Satoh, T.

Arch. Biochem. Biophys. 277, 219-227, 1990
A;Title: Characterization of molecular species of liver microsomal carboxylesterases of

A;Reference number: S09021; MUID:90179180; PMID:2310190

A;Accession: S09021

A;Molecule type: protein

A;Residues: 1-22 <HOS>

A;Cross-references: UNIPARC:UPI00001758FF

C;Superfamily: cholinesterase; cholinesterase homology

C;Keywords: carboxylic ester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
Db 12 KGVVLG 17

RESULT 113

A45140

fatty-acyl-ethyl-ester synthase (EC 3.1.1.67) - rat (fragment)

A;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Dec-2004

C;Accession: A45140

R;Tsujita, T.; Okuda, H.

J. Biol. Chem. 267, 23489-23494, 1992

A;Title: Fatty acid ethyl ester synthase in rat adipose tissue and its relationship to

A;Reference number: A45140; MUID:93054696; PMID:1429692

A;Accession: A45140

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-27 <TSU>

A;Cross-references: UNIPROT:Q9R135; UNIPROT:Q91YG2; UNIPARC:UPI00001758PD

C;Superfamily: cholinesterase homology

C;Keywords: carboxylic ester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
Db 12 KGVVLG 17

RESULT 114

PQ0561

nonstructural protein 3 (clone 9) - hepatitis C virus (fragment)

C;Species: hepatitis C virus

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

C;Accession: PQ0561

R;Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno

Biochem. Biophys. Res. Commun. 181, 279-285, 1991

A;Title: Distribution of plural HCV types in Japan.

A;Reference number: PQ0554; MUID:92068204; PMID:1720309

A;Accession: PQ0561

A;Molecule type: mRNA

A;Residues: 1-41 <KAT>

A;Cross-references: UNIPROT:Q81253; UNIPARC:UPI00000F065B; GB:D10564; GB:D90520; NID:922

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: nonstructural protein

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9

RESULT 115


```
PQ0560
nonstructural protein 3 (clone 8) - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: PQ0560
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Distribution of plural HCV types in Japan.
A:Reference number: PQ0554; MUID:92068204; PMID:1720309
A:Accession: PQ0560
A:Molecule type: mRNA
A:Residues: 1-41 <KAT>
A:Cross-references: UNIPROT:Q81252; UNIPARC:UPI00000F69P4; GB:D10563; GB:D90519; NID:g22
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: nonstructural protein

Query Match          5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
Db 4 GGVLA 9

RESULT 116
PQ0563
nonstructural protein 3 (clone 11) - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: PQ0563
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Distribution of plural HCV types in Japan.
A:Reference number: PQ0554; MUID:92068204; PMID:1720309
A:Accession: PQ0563
A:Molecule type: mRNA
A:Residues: 1-41 <KAT>
A:Cross-references: UNIPROT:Q81248; UNIPARC:UPI00000EE322; GB:D10566; GB:D90522; NID:g22
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: nonstructural protein

Query Match          5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
Db 4 GGVLA 9

RESULT 117
PQ0564
nonstructural protein 3 (clone 4) - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: PQ0564
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Distribution of plural HCV types in Japan.
A:Reference number: PQ0554; MUID:92068204; PMID:1720309
A:Accession: PQ0564
A:Molecule type: mRNA
A:Residues: 1-41 <KAT>
A:Cross-references: UNIPROT:Q81251; UNIPARC:UPI00000F69P4; GB:D10567; GB:D90523; NID:g22
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: nonstructural protein

Query Match          5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
Db 4 GGVLA 9

RESULT 118
PQ0562
nonstructural protein 3 (clone 10) - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: PQ0562
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Distribution of plural HCV types in Japan.
A:Reference number: PQ0554; MUID:92068204; PMID:1720309
A:Accession: PQ0562
A:Molecule type: mRNA
A:Residues: 1-41 <KAT>
A:Cross-references: UNIPROT:Q81247; UNIPARC:UPI00000F65F6; GB:D10565; GB:D90521; NID:g22
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: nonstructural protein

Query Match          5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
Db 4 GGVLA 9

RESULT 119
PQ0565
nonstructural protein 3 (clone 13) - hepatitis C virus (fragment)
C:Species: hepatitis C virus
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: PQ0565
R:Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A:Title: Distribution of plural HCV types in Japan.
A:Reference number: PQ0554; MUID:92068204; PMID:1720309
A:Accession: PQ0565
A:Molecule type: mRNA
A:Residues: 1-41 <KAT>
A:Cross-references: UNIPROT:Q81249; UNIPARC:UPI00000ED385; GB:D10568; GB:D90524; NID:g22
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: nonstructural protein

Query Match          5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
Db 4 GGVLA 9

RESULT 120
Q1BP47
gene 1.4 protein - phage T7
C:Species: phage T7
C>Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 09-Jul-2004
C:Accession: H43002; H43004; S42293; A04416
R:Dunn, J.J.; Thompson, K.
submitted to the Nucleic Acid Sequence Database, September 1982
A:Reference number: A94615
A:Accession: H43002
A:Molecule type: DNA
A:Residues: 1-51 <DUN>
A:Cross-references: UNIPROT:P03791; UNIPARC:UPI00000139383
R:Dunn, J.J.; Studier, F.W.
J. Mol. Biol. 148, 303-330, 1981
A:Title: Nucleotide sequence from the genetic left end of bacteriophage T7 DNA to the be
A:Reference number: A92866; MUID:82078034; PMID:7310871
```

A;Accession: H43004
A;Molecule type: DNA
A;Residues: 1-51 <DU2>
A;Cross-references: UNIPARC:UPI0000139383; GB:V01127; NID:g15498; PIDN:CAA24337.1; PID:g
J;Dunn, J.J.; Studier, F.W.
J. Mol. Biol. 166, 477-535, 1983
A;Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 ge
A;Reference number: 842283; MUID:83241725; PMID:6864790
A;Accession: 842293
A;Molecule type: DNA
A;Residues: 1-51 <DUW>
A;Cross-references: UNIPARC:UPI0000139383; EMBL:V01146; NID:g431187; PIDN:CAA24394.1; PI
C;Genetics:
A;Gene: 1.4
A;Map position: 19.02-19.40
C;Superfamily: phage T7 gene 1.4 protein

Query Match 5.1%; Score 6; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 9 LAALAA 14

RESULT 121
AKQLQ2
adipokinetic hormone II precursor - migratory locus
N;Alternate names: adipokinetic hormone II precursor-related peptide; Lom-AKH-II; peptid
N;Contains: adipokinetic hormone II; adipokinetic hormone-associated peptide II
C;Species: Locusta migratoria (migratory locust)
C;Date: 31-Mar-1988 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
A;Residues: B58652; B25204; B24241; B58653
R;Boigerd, J.; Koolman, F.P.; Pijnenburg, M.A.P.; Hekking, L.H.P.; Oudejans, R.C.H.M.; Va
J. Biol. Chem. 270, 23038-23043, 1995
A;Title: Molecular cloning of three distinct cDNAs, each encoding a different adipokinetic
okinetic hormone precursor genes during flight activity.
A;Reference number: A58652; MUID:96032738; PMID:7559443
A;Accession: B58652
A;Molecule type: mRNA
A;Residues: 1-61 <BOG>
A;Cross-references: UNIPROT:P08379; UNIPARC:UPI000012576A; GB:X86800; NID:g1085062; PIDN
R;Siebert, K.; Morgan, P.; Mordue, W.
Biol. Chem. Hoppe-Seyler 366, 723-727, 1985
A;Title: Primary structures of locust adipokinetic hormones II.
A;Reference number: A90692; MUID:86050918; PMID:4063072
A;Accession: B25204
A;Molecule type: protein
A;Residues: 23-30 <SIE>
A;Cross-references: UNIPARC:UPI00001734F2
R;Gade, G.; Goldsworthy, G.J.; Schaffer, M.H.; Cook, J.C.; Rinehart Jr., K.L.
Biochem. Biophys. Res. Commun. 134, 723-730, 1986
A;Title: Sequence analyses of adipokinetic hormones II from corpora cardiaca of Schistoc
A;Reference number: A24241; MUID:86130555; PMID:3947348
A;Accession: B24241
A;Molecule type: protein
A;Residues: 'E', 24-30 <GAD>
A;Cross-references: UNIPARC:UPI00001734F3
R;Hietter, H.; Luu, B.; Goltzene, F.; Zachary, D.; Hoffmann, J.; Van Dorsselaer, A.
Eur. J. Biochem. 182, 77-84, 1989
A;Title: Isolation and structure of two novel 6-kDa dimeric peptides from the corpora ca
A;Reference number: A58653; MUID:89276392; PMID:2731552
A;Accession: B58653
A;Molecule type: protein
A;Residues: 34-61 <HIE>
A;Cross-references: UNIPARC:UPI00001734F4
A;Experimental source: corpora cardiaca
C;Complex: forms disulfide linked homodimers; disulfide linked heterodimers form between
C;Function:
A;Description: hormone, released from the corpora cardiaca after the beginning of flight
s an energy source
C;Superfamily: adipokinetic hormone

C;Keywords: amidated carboxyl end; corpora cardiaca; heterodimer; homodimer; hormone; ne
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-30/Product: adipokinetic hormone II #status experimental <MAT1>
F;34-61/Product: adipokinetic hormone-associated peptide II #status predicted <MAT2>
F;23/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experimen
F;30/Modified site: amidated carboxyl end (Trp) (amide in mature form from following gly
F;59/Binding site: glutathione (Cys) (covalent) (partial) #status experimental
F;59/Disulfide bonds: interchain (partial) #status experimental
F;59/Disulfide bonds: interchain (to adipokinetic hormone-associated peptide I 61) (part

Query Match 5.1%; Score 6; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
|||||
Db 15 VLAALA 20

RESULT 122
D91257
hypothetical protein ECs5028 [imported] - Escherichia coli (strain O157:H7, substrain RI.
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: D91257
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: D91257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-69 <HAY>
A;Cross-references: UNIPROT:P32691; UNIPARC:UPI0000113221; GB:BA000007; PIDN:BA838451.1;
C;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs5028

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGV 86
|||||
Db 12 QPKGV 17

RESULT 123
H86097
hypothetical protein yjbj [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: H86097
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Killer, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: H86097
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-69 <STO>
A;Cross-references: UNIPROT:P32691; UNIPARC:UPI0000113221; GB:AE005174; NID:g12518994; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: yjbj

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGV 86

```

Db      12 QFKGV 17
|||||
RESULT 124
hypothetical protein b4045 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: D65212
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: D65212
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-69 <BLAT>
A:Cross-references: UNIPROT:P32691; UNIPARC:UPI0000113221; GB:AE000478; GB:U00096; NID:9
A:Experimental source: strain K-12, substrain MG1655

Query Match      5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QFKGV 86
|||||
Db 12 QFKGV 17

RESULT 125
E84364
H+-transporting ATP synthase subunit K [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84364
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: E84364
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-71 <STO>
A:Cross-references: UNIPROT:Q9HND9; UNIPARC:UPI0000063A65; GB:AE004437; NID:g10581558; F
C:Genetics:
A:Gene: atpK

Query Match      5.1%; Score 6; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 16 LAALAA 21

RESULT 126
G70534
hypothetical protein Rv0660c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: G70534
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

```

```

A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70534
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-81 <COL>
A:Cross-references: UNIPROT:O06779; UNIPARC:UPI000000C14D7; GB:Z95972; GB:AL123456; NID:
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0660c

```

```

Query Match      5.1%; Score 6; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 37 LAALAA 42

```

```

RESULT 127
AC1829
hypothetical protein asr0179 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC1829
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A:Title: Complete genomic sequence of the filamentous Nitrogen-fixing Cyanobacterium An
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1829
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-81 <KUR>
A:Cross-references: UNIPROT:Q8Z0B9; UNIPARC:UPI000000CDC67; GB:BA000019; PIDN:BA877703.1,
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asr0179

```

```

Query Match      5.1%; Score 6; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 46 PAIVPD 51
|||||
Db 66 PAIVPD 71

```

```

RESULT 128
B84210
hypothetical protein Vng0516h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84210
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: B84210
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <STO>
A:Cross-references: UNIPROT:Q9HRW2; UNIPARC:UPI000006365C; GB:AE004437; NID:g10580118;
C:Genetics:
A:Gene: VNG0516H

```

```

Query Match      5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 17 LGGVLA 22
 |||||
 Db 31 LGGVLA 36

RESULT 129

T35070

probable membrane protein - Streptomyces coelicolor

C/Species: Streptomyces coelicolor

C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C/Accession: T35070

R/Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, July 1999

A/Reference number: Z21567

A/Accession: T35070

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-83 <SEE>

A/Cross-references: UNIPROT:Q9S2U4; UNIPARC:UPI00000DB2A6; EMBL:AL096884; PIDN:CAB51433.

A/Experimental source: strain A3(2)

C/Genetics:

A/Gene: SCORDB:SC4G6.10C

Query Match 5.1%; Score 6; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20

|||||

Db 45 VLLGGV 50

RESULT 130

T44917

proteolipid precursor [imported] - Halobacterium salinarum

C/Species: Halobacterium salinarum

C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

C/Accession: T44917

R/Kunio, I.; Satohshi, W.; Ken-ichiro, S.; Yasuo, M.

submitted to the EMBL Data Library, August 1996

A/Description: DCCD binding proteolipid from extremely halophilic archaeson Halobacterium

A/Reference number: Z22866

A/Accession: T44917

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-89 <KUN>

A/Cross-references: UNIPROT:Q48302; UNIPARC:UPI00000625A7; EMBL:D86915; PIDN:BAAL13179.1

C/Genetics:

A/Gene: atp P

F,i-12/Domain: signal sequence #status predicted <SIG>

F,i3-89/Product: proteolipid #status predicted <MAT>

Query Match 5.1%; Score 6; DB 2; Length 89;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

|||||

Db 34 LAALAA 39

RESULT 131

F70630

hypothetical protein Rv0424c - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C/Accession: F70630

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

R/Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: F70630

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-91 <COL>

A/Cross-references: UNIPROT:P96270; UNIPARC:UPI00000C14BD; GB:Z84724; GB:AL123456; NID:9

A/Experimental source: strain H37RV

C/Genetics:

A/Gene: RV0424c

Query Match 5.1%; Score 6; DB 2; Length 91;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

|||||

Db 49 LAALAA 54

RESULT 132

S41724

hypothetical protein - Saccharopolyspora erythraea

C/Species: Saccharopolyspora erythraea

C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C/Accession: S41724

R/Brown, D.P.; Idler, K.B.; Backer, D.M.; Donadio, S.; Katz, L.

Mol. Gen. Genet. 242, 185-193, 1994

A/Title: Characterization of the genes and attachment sites for site-specific integration

A/Reference number: S41722; MUID:94211208; PMID:8159169

A/Accession: S41724

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-92 <BRO>

A/Cross-references: UNIPROT:Q54078; UNIPARC:UPI00000B16D6; EMBL:L11597; NID:G404788; PID

Query Match 5.1%; Score 6; DB 2; Length 92;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

|||||

Db 13 VLAALA 18

RESULT 133

D90249

hypothetical protein SS06904 [imported] - Sulfolobus solfataricus

C/Species: Sulfolobus solfataricus

C/Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C/Accession: D90249

R/She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A/Description: Sulfolobus solfataricus complete genome.

A/Reference number: A99139

A/Accession: D90249

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-94 <KUR>

A/Cross-references: UNIPROT:Q97ZE1; UNIPARC:UPI0000064301; GB:AB006641; NID:G13814161; P

C/Genetics:

A/Gene: SS06904

Query Match 5.1%; Score 6; DB 2; Length 94;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGVK 86

|||||

Db 17 QFKGVK 22

RESULT 134

P83404
hypothetical protein PA1929 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83404
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: F83404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <STO>
A:Cross-references: UNIPROT:Q9I2H5; UNIPARC:UPI00000C5511; GB:AE004619; GB:AE004091; NID:10984043
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA1929

Query Match 5.1%; Score 6; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 22 LAALAA 27

RESULT 135

S44892
ZK112.4 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S44892
R:Du, Z.
submitted to the EMBL Data Library, May 1993
A:Description: Sequence of the C. elegans cosmid ZK112.
A:Reference number: S44616
A:Accession: S44892
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <DUZ>
A:Cross-references: UNIPROT:P34613; UNIPARC:UPI000013BC79; EMBL:L14324; NID:g289740; PID:10984043
C:Superfamily: Caenorhabditis elegans ZK112.4 protein

Query Match 5.1%; Score 6; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIELGG 44
|||||
DB 49 HIELGG 54

RESULT 136

T32603
hypothetical protein F09D12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32603
R:Delehaanty, A.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F09D12.
A:Reference number: T32198
A:Accession: T32603
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102
A:Cross-references: UNIPROT:O44471; UNIPARC:UPI000007F5B0; EMBL:AF038610; PIDN:AAB92033.
A:Experimental source: strain Bristol N2; clone F09D12

C:Genetics:

A:Gene: CESP:F09D12.1
A:Map position: 4
A:Introns: 27/1

Query Match 5.1%; Score 6; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 6 VLAALA 11

RESULT 137

T31879
hypothetical protein F41E6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31879
R:Sammons, L.; Murray, J.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid F41E6.
A:Reference number: Z21095
A:Accession: T31879
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-102 <SAM>
A:Cross-references: UNIPROT:O16462; UNIPARC:UPI000007DDAD; EMBL:AF016448; PIDN:AAB65952
A:Experimental source: strain Bristol N2; clone F41E6
C:Genetics:
A:Gene: CESP:F41E6.2
A:Map position: 5
A:Introns: 27/1

Query Match 5.1%; Score 6; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 6 VLAALA 11

RESULT 138

A12693
conserved hypothetical protein Atu0953 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 31-Dec-2004
C:Accession: A12693
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, J.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: A12693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-102 <KUR>
A:Cross-references: UNIPROT:Q8UGT2; UNIPARC:UPI0000164565; GB:AE008688; PIDN:AAL41967.1.
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0953
A:Map position: circular chromosome
C:Superfamily: wound-responsive protein

Query Match 5.1%; Score 6; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
Db 61 AALAA 66

RESULT 139

H87548
hypothetical protein CC2417 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: H87548
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: H87548
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-104 <STO>
A/Cross-references: UNIPROT:Q9A5N0; UNIPARC:UPI00000C770C; GB:AE005673; NID:g13423959; F
C/Genetics:
A/Gene: CC2417

Query Match 5.1%; Score 6; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 13 LAALAA 18

RESULT 140

TPXL2
insulin II precursor - African clawed frog
C/Species: Xenopus laevis (African clawed frog)
C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C/Accession: B33847; S13537
R/Shuldiner, A.R.; Phillips, S.; Roberts Jr., C.T.; LeRoith, D.; Roth, J.
J. Biol. Chem. 264, 9428-9432, 1989
A/Title: Xenopus laevis contains two nonallelic preproinsulin genes. cDNA cloning and ev
A/Reference number: A33847; MUID:89255444; PMID:2722842
A/Accession: B33847
A/Molecule type: mRNA
A/Residues: 1-106 <SHU1>
A/Cross-references: UNIPROT:P12707; UNIPARC:UPI000012D6A3; GB:M24442; GB:J04796; NID:g21
R/Shuldiner, A.R.; Bennett, C.; Robinson, E.A.; Roth, J.
Endocrinology 125, 469-477, 1989
A/Title: Isolation and characterization of two different insulins from an amphibian, Xen
A/Reference number: 807199; MUID:89289601; PMID:2661211
A/Accession: S13537

A/Molecule type: protein
A/Residues: 24-53;86-106 <SHU2>
A/Cross-references: UNIPARC:UPI000017355F; UNIPARC:UPI0000173560
C/Superfamily: insulin
C/Keywords: hormone; pancreas
F/1-23/Domain: signal sequence #status predicted <SIG>
F/24-53/Domain: insulin chain B #status experimental <BCH>
F/24-53;86-106/Product: insulin #status experimental <MAT>
F/56-83/Domain: connecting peptide #status predicted <CPEP>
F/86-106/Domain: insulin chain A #status experimental <ACH>
F/30-92,42-105,91-96/Disulfide bonds: #status predicted

Query Match 5.1%; Score 6; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 IEQAQV 77
Db 57 IEQAQV 62

RESULT 141

A43301
ribosomal protein L36a.e, cytosolic - yeast (Candida maltosa)
N/Alternate names: ribosomal protein YL41
C/Species: Candida maltosa
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: A43301
R/Kawai, S.; Muroa, S.; Mochizuki, M.; Shibuya, I.; Yano, K.; Takagi, M.
J. Bacteriol. 174, 254-262, 1992
A/Title: Drastic alteration of cycloheximide sensitivity by substitution of one amino ac
A/Reference number: A43301; MUID:92104971; PMID:1729213
A/Accession: A43301
A/Molecule type: DNA
A/Residues: 1-106 <KAW>
A/Cross-references: UNIPROT:P27074; UNIPARC:UPI00001698BF; GB:D10577; GB:D90488; NID:g21
A/Note: sequence extracted from NCBI backbone (NCBIN:75358, NCBIP:75359)
C/Superfamily: rat ribosomal protein L36a
C/Keywords: protein biosynthesis; ribosome

Query Match 5.1%; Score 6; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELGG 44
Db 90 HIELGG 95

RESULT 142

D43301
ribosomal protein L36a.e, cytosolic - yeast (Candida tropicalis)
N/Alternate names: ribosomal protein YL41
C/Species: Candida tropicalis
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: D43301
R/Kawai, S.; Muroa, S.; Mochizuki, M.; Shibuya, I.; Yano, K.; Takagi, M.
J. Bacteriol. 174, 254-262, 1992
A/Title: Drastic alteration of cycloheximide sensitivity by substitution of one amino ac
A/Reference number: A43301; MUID:92104971; PMID:1729213
A/Contents: N7Y1
A/Accession: D43301
A/Molecule type: DNA
A/Residues: 1-106 <KAW>
A/Cross-references: UNIPROT:P27075; UNIPARC:UPI00001698CF; GB:D10581; GB:D90492; NID:g21
A/Note: sequence extracted from NCBI backbone (NCBIN:75364, NCBIP:75365)
C/Superfamily: rat ribosomal protein L36a
C/Keywords: protein biosynthesis; ribosome

Query Match 5.1%; Score 6; DB 2; Length 106;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELGG 44
Db 90 HIELGG 95

RESULT 143

C87494
conserved hypothetical protein CC1976 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: C87494
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: C87494
A/Status: preliminary
A/Molecule type: DNA

A:Residues: 1-108 <STO>
A:Cross-references: UNIPROT:Q9A6V7; UNIPARC:UPI000013B4D3; GB:AE005673; NID:g13423441; F
C:Genetics:
A:Gene: CC1976
C:Superfamily: hypothetical protein b1582

Query Match 5.1%; Score 6; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 8 VLAALA 13

RESULT 144
C72392
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: C72392
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: C72392
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <ARN>
A:Cross-references: UNIPROT:Q9WTF2; UNIPARC:UPI00000C1444; GB:AE001713; GB:AE000512; NID
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0316

Query Match 5.1%; Score 6; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 37 VLAALA 42

RESULT 145
A70596
probable ACETYLTRANSFERASE - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: A70596
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70596
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-110 <COL>
A:Cross-references: UNIPROT:Q05850; UNIPARC:UPI00000D10F5; GB:Z95120; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3216

Query Match 5.1%; Score 6; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

A:Residues: 1-108 <STO>
A:Cross-references: UNIPROT:Q9A6V7; UNIPARC:UPI000013B4D3; GB:AE005673; NID:g13423441; F
C:Genetics:
A:Gene: CC1976
C:Superfamily: hypothetical protein b1582

Query Match 5.1%; Score 6; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 8 VLAALA 13

RESULT 144
C72392
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: C72392
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: C72392
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <ARN>
A:Cross-references: UNIPROT:Q9WTF2; UNIPARC:UPI00000C1444; GB:AE001713; GB:AE000512; NID
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0316

Query Match 5.1%; Score 6; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 37 VLAALA 42

RESULT 145
A70596
probable ACETYLTRANSFERASE - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: A70596
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70596
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-110 <COL>
A:Cross-references: UNIPROT:Q05850; UNIPARC:UPI00000D10F5; GB:Z95120; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3216

Query Match 5.1%; Score 6; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

Db 103 LAALAA 108

RESULT 146

D75445

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: D75445

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036996; PMID:10567266

A:Accession: D75445

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <WHI>

A:Cross-references: UNIPROT:Q9RVI6; UNIPARC:UPI00000C1873; GB:AE001955; GB:AE000513; NID

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1043

A:Map position: 1

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 112;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24

|||||

DB 85 GVLAAAL 90

RESULT 147

I40680

acetyl-CoA C-acetyltransferase (EC 2.3.1.9) - Clostridium difficile (fragment)

N:Alternate names: thiolase

C:Species: Clostridium difficile

C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004

C:Accession: I40680; S49139

R:Mullany, P.; Clayton, C.L.; Pallen, M.J.; Slone, R.; al-Saleh, A.; Tabagchali, S.

FEMS Microbiol. Lett. 124, 61-67, 1994

A:Title: Genes encoding homologues of three consecutive enzymes in the butyrate/butanol-

A:Reference number: I40678; MUID:95095030; PMID:8001771

A:Accession: I40680

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-113 <RES>

A:Cross-references: UNIPROT:P45362; UNIPARC:UPI000016EACF; EMBL:X79899; NID:g509743; PID

R:Mullany, P.; Clayton, C.L.; Pallen, M.J.; Al-Salem, A.; Tabagchali, S.

submitted to the EMBL Data Library, April 1994

A:Description: Genes encoding homologues of three consecutive enzymes in the butyrate/bu

A:Reference number: S49138

A:Accession: S49139

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-112 <MUL>

A:Cross-references: UNIPARC:UPI0000136EAF; EMBL:X79899

C:Genetics:

A:Gene: thi

C:Superfamily: acetyl-CoA acetyltransferase

C:Keywords: acyltransferase; coenzyme A

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 113;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLT 21

|||||

DB 57 LLGGVLT 62

RESULT 148

D75507
 A:Title: hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C/Accession: D75507
 R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A/Accession: D75507
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <WHI>
 A/Cross-references: UNIPROT:Q9RWY9; UNIPARC:UPI00000C1785; GB:AE001911; GB:AE000513; NID
 A/Experimental source: strain R1
 C/Genetics:
 A:Gene: DR0526
 A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
 |||||
 Db 3 GVLAAL 8

RESULT 149

E95174
 A:Title: bacteriocin transport accessory protein SPI499 [imported] - Streptococcus pneumoniae (str
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C/Accession: E95174
 R/Tattellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfle,
 neon, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Lofthus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
 A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A/Accession: E95174
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <KUR>
 A/Cross-references: UNIPROT:Q97PU4; UNIPARC:Q8DP51; UNIPARC:UPI000005188E; GB:AE005672;
 A/Experimental source: strain TIGR4
 C/Genetics:
 A:Gene: SPI499

Query Match 5.1%; Score 6; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVTT 11
 |||||
 Db 10 DLEVTT 15

RESULT 150

G98040
 A:Title: bacteriocin transport accessory protein [imported] - Streptococcus pneumoniae (strain R6)
 C:Species: Streptococcus pneumoniae
 C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C/Accession: G98040
 R/Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234

A/Accession: G98040
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <KUR>
 A/Cross-references: UNIPROT:Q97PU4; UNIPARC:UPI000005188E; GB:AE007317;
 C/Genetics:
 A:Gene: bta

Query Match 5.1%; Score 6; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVTT 11
 |||||
 Db 10 DLEVTT 15

RESULT 151

F95862
 A:Title: conserved hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magapla
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C/Accession: F95862
 R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A/Accession: F95862
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-117 <KUR>
 A/Cross-references: UNIPROT:Q9ZWZ6; UNIPARC:UPI000000CB432; GB:AL591985; PIDN:CAC48566.1;
 A/Experimental source: strain 1021, megaplasmid pSymB
 R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.;
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 C/Contents: annotation
 C/Genetics:
 A:Gene: SMB20166
 A:Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 117;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 |||||
 Db 6 LAALAA 11

RESULT 152

F89756
 A:Title: protein T2J87.4 [imported] - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C/Accession: F89756
 R/anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
 A:Reference number: A75000; MUID:99069613; PMID:9851916
 A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A/Accession: F89756
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-117 <STO>


```
A:Cross-references: UNIPROT:O17341; UNIPARC:UPI0000076E45; GB:chr_X; PIDN:AAB71255.1; PI
C:Genetics:
A:Map position: X

Query Match      5.1%; Score 6; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 MSADLE 8
    |||||
DB 94 MSADLE 99

RESULT 153
D72520
hypothetical protein APE2137 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: D72520
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72520
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-119 <KAW>
A:Cross-references: UNIPROT:Q9YA02; UNIPARC:UPI000005E218; DDBJ:AF000063; NID:G5105654;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2137
C:Superfamily: Aeropyrum pernix hypothetical protein APE2137

Query Match      5.1%; Score 6; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 TTSTWV 15
    |||||
DB 26 TTSTWV 31

RESULT 154
B72628
hypothetical protein APE1482 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B72628
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72628
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <KAW>
A:Cross-references: UNIPROT:Q9YBW8; UNIPARC:UPI000005DF7D; DDBJ:AF000061; NID:G5104821;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1482
C:Superfamily: Aeropyrum pernix hypothetical protein APE1482

Query Match      5.1%; Score 6; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GKPAIV 49
    |||||
DB 22 GKPAIV 27
```

RESULT 155

AB2389

hypothetical protein all4666 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AB2389

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, A.;

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An-

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AB2389

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-124 <KUR>

A:Cross-references: UNIPROT:Q8YNA1; UNIPARC:UPI000000CEBEB; GB:BA000019; PIDN:BAB76365.1;

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4666

Query Match

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 ATQOQA 98

DB 9 ATQOQA 14

RESULT 156

S01397

H+-transporting two-sector ATPase (EC 3.6.3.14) chain I - thermophilic bacterium PS-3

C:Species: thermophilic bacterium PS-3

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 27-Oct-2003

C:Accession: S01397

R:Ohta, S.; Yohda, M.; Ishizuka, M.; Hirata, H.; Hamamoto, T.; Otawara-Hamamoto, Y.; Mat

Biochim. Biophys. Acta 933, 141-155, 1998

A:Title: Sequence and over-expression of subunits of adenosine triphosphate synthase in

A:Reference number: S01397; MUID:88163879; PMID:2894854

A:Accession: S01397

A:Molecule type: DNA

A:Residues: 1-127 <OHT>

A:Cross-references: UNIPARC:UPI0000126610; EMBL:X07804; NID:G45808; PIDN:CAA30647.1; PI

A:Superfamily: F0F1-ATP synthase auxiliary protein i, gram positive type

C:Keywords: ATP biosynthesis; hydrolase

Query Match

Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26

DB 80 LAALAA 85

RESULT 157

A54670

RNA polymerase II transcription cofactor p15 - human

N:Alternate names: DNA binding protein PC4

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C:Accession: A54670; A54669

R:Kretschmar, M.; Kaiser, K.; Lottspeich, F.; Meisterernst, M.

Cell 78, 525-534, 1994

A:Title: A novel mediator of class II gene transcription with homology to viral immediat

A:Reference number: A54670; MUID:94340741; PMID:8062392

A:Accession: A54670

A:Molecule type: mRNA

A:Residues: 1-127 <KRE>

A:Cross-references: UNIPROT:P53999; UNIPARC:UPI0000044938; GB:X79805; NID:G619160; PIDN

R:Ge, H.; Roeder, R.G.

Cell 78, 513-523, 1994
A:Title: Purification, cloning, and characterization of a human coactivator, PC4, that m
A:Reference number: A54669; MUID:94340740; PMID:8062391
A:Accession: A54669
A:Molecule type: mRNA
A:Residues: 1-127 <GBA>
A:Cross-references: UNIPARC:UPI0000044938; GB:U12979; NID:G531394; PIDN:AAA20980.1; PID:
A:Note: parts of this sequence, including the amino end of the mature protein, were con
C:Genetics:
A:Gene: GDB:P15
A:Cross-references: GDB:453302
C:Superfamily: Caenorhabditis elegans hypothetical protein T13F2.2
C:Keywords: DNA binding; nucleus; phosphoprotein
F2:127/Product: RNA polymerase II transcription cofactor p15 #status experimental <MAY>

Query Match 5.1%; Score 6; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PKGKVL 87
Db 77 PKGKVL 82
|||||

RESULT 158
C70971
hypothetical protein Rv3364c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70971
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID:9829987; PMID:9634230
A: Accession: C70971
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-130 <COL>
A: Cross-references: UNIPROT:O50393; UNIPARC:UPI00000D1076; GB:AL009198; GB:AL123456; NID:
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: Rv3364c
C: Superfamily: Streptomyces coelicolor hypothetical protein SC1C2.20c

Query Match 5.1%; Score 6; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 97 LAALAA 102
|||||

RESULT 159
F97475
Id894 (AF322013) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 31-Dec-2004
C:Accession: F97475
R: Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A: Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A: Reference number: A97359; MUID:21608551; PMID:11743194
A: Accession: F97475
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-134 <KUR>
A: Cross-references: UNIPROT:Q8UGT2; UNIPARC:UPI00000D19B9; GB:AB007869; PIDN:AAK86759.1;
C: Genetics:

A:Gene: AGR_C_1741
A:Map position: circular chromosome
C:Superfamily: wound-responsive protein

Query Match 5.1%; Score 6; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
Db 93 AALAA 98
|||||

RESULT 160
C70660
hypothetical protein Rv2551c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70660
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID:9829987; PMID:9634230
A: Accession: C70660
A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-139 <COL>
A: Cross-references: UNIPROT:P95002; UNIPARC:UPI00000D1049; GB:Z83863; GB:AL123456; NID:
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: Rv2551c
C: Superfamily: Mycobacterium tuberculosis hypothetical protein Rv2551c

Query Match 5.1%; Score 6; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 95 LAALAA 100
|||||

RESULT 161
H81886
hypothetical protein NMA1191 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: H81886
R: Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A: Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A: Reference number: A81775; MUID:20222556; PMID:10761919
A: Accession: H81886
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-139 <PAR>
A: Cross-references: UNIPROT:Q9JUQ5; UNIPARC:UPI00000C4B39; GB:AL162755; GB:AL157959; NID:
A: Experimental source: serogroup A, strain Z2491
C: Genetics:
A: Gene: NMA1191
C: Superfamily: Neisseria meningitidis hypothetical protein NMA1191

Query Match 5.1%; Score 6; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 QQAVIE 101
Db 39 QQAVIE 44
|||||

RESULT 162

S45924
probable membrane protein YBR064w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein YBR0612
C:Species: *Saccharomyces cerevisiae*
C:Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
A:Accession: S45924
R:Domdey, H.; Gassenhuber, H.; Obermaier, B.; Piravandi, E.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45816
A:Accession: S45924
A:Molecule type: DNA
A:Residues: 1-142 <DOM>
A:Cross-references: UNIPROT:P38240; UNIPARC:UPI000013A3CD; EMBL:Z35932; NID:G536306; PID

A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YBR064w
A:Cross-references: SGD:S0000268
A:Map position: 2R

C:Superfamily: *Saccharomyces cerevisiae* probable membrane protein YBR064w
C:Keywords: transmembrane protein
P:13-30/Domain: transmembrane #status predicted <TM1>
P:48-64/Domain: transmembrane #status predicted <TM2>
P:98-127/Domain: transmembrane #status predicted <TM3>

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 142;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 101 EPIVTT 106

|||||

DB 68 EPIVTT 73

RESULT 163

AF1517
hypothetical protein lin0678 [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
A:Accession: AF1517
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1517
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-142 <GLA>
A:Cross-references: UNIPROT:Q92DY3; UNIPARC:UPI00000CC2PB; GB:AL592022; PIDN:CAC95910.1;

A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0678

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 142;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLT 21

|||||

DB 70 LGGVLT 75

RESULT 164

H84292
hypothetical protein Vng1382h [imported] - *Halobacterium* sp. NRC-1
C:Species: *Halobacterium* sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: H84292

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laskey,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: H84292

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <STO>

A:Cross-references: UNIPROT:Q9HQ09; UNIPARC:UPI0000063890; GB:AE004437; NID:G10580890; E

C:Genetics:

A:Gene: VNG1382H

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 144;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23

|||||

DB 7 GGVLA 12

RESULT 165

D70196

hypothetical protein BB0773 - Lyme disease spirochete

C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 09-Jul-2004

C:Accession: D70196

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Whit

son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,

; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943; PMID:9403685

A:Accession: D70196

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-144 <KLS>

A:Cross-references: UNIPROT:O51714; UNIPARC:UPI00000575E7; GB:AE001177; GB:AE000783; NID

A:Experimental source: strain B31

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 144;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 VIEPIV 104

|||||

DB 77 VIEPIV 82

RESULT 166

S63582

foli protein - *Neisseria gonorrhoeae* (strain MS11)

C:Species: *Neisseria gonorrhoeae*

A:Variety: strain MS11

C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

A:Accession: S63582

R:Fusenegeger, M.; Meyer, T.F.

Mol. Gen. Genet. 250, 277-285, 1996

A:Title: Cloning and characterization of the *Neisseria gonorrhoeae* MS11 folC gene.

A:Reference number: S63582; MUID:96180644; PMID:860142

A:Accession: S63582

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-146 <FUS>

A:Cross-references: UNIPROT:Q50989; UNIPARC:UPI00000BF01D; GB:Z68205; NID:G1237075; PID

C:Genetics:

A:Gene: folI

C:Superfamily: *Neisseria gonorrhoeae* folI protein

Query Match 5.1%; Score 6; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 |||||

Db 42 LAALAA 47
 |||||

RESULT 167
 C70314
 thiol-disulfide interchange protein tlpA [imported] - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
 C/Accession: C70314
 R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

Nature 392, 353-358, 1998
 A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A/Reference number: A70300; MUID:98196666; PMID:9537320
 A/Accession: C70314
 A/Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-146 <AQP>
 A/Cross-references: UNIPROT:O66542; UNIPARC:UPI0000056299; GB:AE000675; NID:G2982863; PI
 A/Experimental source: strain VF5
 C/Genetics:
 A/Gene: tlpA

Query Match 5.1%; Score 6; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 FKGVKL 87
 |||||

Db 58 FKGVKL 53
 |||||

RESULT 168
 A83374
 hypothetical protein PA2169 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C/Accession: A83374
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: A83374
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-150 <STO>
 A/Cross-references: UNIPROT:Q91IU6; UNIPARC:UPI00000C55E5; GB:AE004644; GB:AE004091; NID
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA2169

Query Match 5.1%; Score 6; DB 2; Length 150;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKP 46
 |||||

Db 53 ELGGKP 58
 |||||

RESULT 169
 AF0887
 conserved hypothetical protein STY3341 [imported] - Salmonella enterica subsp. enterica
 C/Species: Salmonella enterica subsp. enterica serovar Typhi

A/Note: this species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 31-Dec-2004
 C/Accession: AF0887
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AF0887
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-151 <PAR>
 A/Cross-references: UNIPARC:UPI00000CDB9F; GB:AL513382; PIDN:CAD02998.1; PID:G16504244;
 C/Genetics:
 A/Gene: STY3341
 C/Superfamily: histidine triad hydrolase

Query Match 5.1%; Score 6; DB 2; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 |||||

Db 141 LAALAA 146
 |||||

RESULT 170
 F84070
 spore coat polysaccharide synthesis (dTDP-4-dehydrorhamnose reductase) spsk [imported]
 C/Species: Bacillus halodurans
 C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C/Accession: F84070
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A/Reference number: A83650; MUID:20512582; PMID:11058132
 A/Accession: F84070
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-151 <STO>
 A/Cross-references: UNIPROT:Q9K7J5; UNIPARC:UPI00000C419A; GB:AP001518; GB:BA000004; NID
 A/Experimental source: strain C-125
 C/Genetics:
 A/Gene: spsk
 C/Superfamily: dTDP-4-dehydrorhamnose 3,5-epimerase

Query Match 5.1%; Score 6; DB 2; Length 151;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPA 47
 |||||

Db 110 LGGKPA 115
 |||||

RESULT 171
 G72647
 hypothetical protein APE0613 - Aeropyrum pernix (strain K1)
 C/Species: Aeropyrum pernix
 C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C/Accession: G72647
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
 DNA Res. 6, 83-101, 1999
 A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
 A/Reference number: A72450; MUID:99310339; PMID:10382966
 A/Accession: G72647
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-151 <KAW>
 A/Cross-references: UNIPROT:Q9YEG3; UNIPARC:UPI000005DBFC; DDBJ:AP000060; NID:G5104188;

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0613

C:Superfamily: Aeropyrum pernix hypothetical protein APE0613

Query Match 5.1%; Score 6; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20

DB 83 VLLGGV 88

RESULT 172

H64243

galactoside O-acetyltransferase (EC 2.3.1.18) - Mycoplasma genitalium

C:Species: Mycoplasma genitalium

C:Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 31-Dec-2004

C:Accession: H64243

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;

M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.

C.A.; Venter, J.C.

Science 270, 397-403, 1995

A:Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; MUID:96026346; PMID:7569993

A:Accession: H64243

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-152 <TIG>

A:Cross-references: UNIPROT:P47636; UNIPARC:UPI000013973C; GB:U39725; GB:L43967; NID:g10

A:Experimental source: strain G-37

C:Genetics:

A:Genetic code: SGC3

C:Superfamily: sugar-phosphate isomerase, RpiB/LacA/LacB types

C:Keywords: acyltransferase; coenzyme A

Query Match 5.1%; Score 6; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24

DB 84 GVLAAL 89

RESULT 173

P85642

hypothetical protein Z1481 [imported] - Escherichia coli (strain O157:H7, substrain EDL9

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: P85642

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: P85642

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 <STO>

A:Cross-references: UNIPROT:Q8XAX9; UNIPARC:UPI00000D0665; GB:AE005174; NID:g12514334; E

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: Z1481

Query Match 5.1%; Score 6; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

DB 79 VLAALA 84

RESULT 174

A90782

hypothetical protein ECs1225 [imported] - Escherichia coli (strain O157:H7, substrain R

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C:Accession: A90782

R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A:Reference number: A9629; MUID:21156231; PMID:11258796

A:Accession: A90782

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 <HAV>

A:Cross-references: UNIPROT:Q8XAX9; UNIPARC:UPI00000D0665; GB:BA000007; PIDN:BA034648.1

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: ECs1225

Query Match 5.1%; Score 6; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

DB 79 VLAALA 84

RESULT 175

S77145

ybeB protein homolog iojap - Synechocystis sp. (strain PCC 6803)

N:Alternate names: protein slr1886

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C:Accession: S77145

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.

O.K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocysti

s.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S77145

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-154 <KAN>

A:Cross-references: UNIPROT:P73658; UNIPARC:UPI00000D3409; EMBL:D90908; GB:AB001339; NII

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

C:Superfamily: Escherichia coli ybeB protein

Query Match 5.1%; Score 6; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 KLEAFW 115

DB 124 KLEAFW 129

RESULT 176

E96728

hypothetical protein F24J13.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E96728

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: E96728
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-154 <STO>
A;Cross-references: UNIPROT:Q9CAL7; UNIPARC:UPI00000AA24D; GB:AE005173; NID:g6175139; P
C;Genetics:
A;Gene: F24J13.4
A;Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 154;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 48 VLAALA 53

RESULT 177
G87445
hypothetical protein CC1584 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87445
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87445
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-155 <STO>
A;Cross-references: UNIPROT:Q9AY75; UNIPARC:UPI00000C742B; GB:AE005673; NID:g13422977; P
C;Genetics:
A;Gene: CC1584
C;Superfamily: peptidylprolyl isomerase; cyclophilin homology

Query Match 5.1%; Score 6; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MSADLE 8
| | | | |
Db 1 MSADLE 6

RESULT 178
F83541
probable glutathione peroxidase PA0838 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83541
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bu
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83541
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-160 <STO>
A;Cross-references: UNIPROT:Q915A2; UNIPARC:UPI00000C5181; GB:AE004518; GB:AE004091; NID

A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0838
C;Superfamily: glutathione peroxidase

Query Match 5.1%; Score 6; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 FKGVKL 87
| | | | |
Db 22 FKGVKL 27

RESULT 179
S36491
E6 protein - human papillomavirus type 25
C;Species: human papillomavirus type 25
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S36491
R;Deilus, H.; Hofmann, B.
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36491
A;Molecule type: DNA
A;Residues: 1-161 <DBL>
A;Cross-references: UNIPROT:P28833; UNIPARC:UPI00001383C4; EMBL:X74471; NID:g396948; PID
C;Superfamily: papillomavirus E6 protein
C;Keywords: DNA binding; early protein; nucleus; zinc finger

Query Match 5.1%; Score 6; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 EQAQVI 78
| | | | |
Db 14 EQAQVI 19

RESULT 180
YAZ051
antigen 5.1 precursor - malaria parasite (Plasmodium falciparum)
C;Species: Plasmodium falciparum
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004
C;Accession: A23052
R;Hope, I.A.; Mackay, M.; Hyde, J.E.; Goman, M.; Scaife, J.
Nucleic Acids Res. 13, 369-379, 1985
A;Title: The gene for an exported antigen of the malaria parasite Plasmodium falciparum
A;Reference number: A23052; MUID:85215483; PMID:2582354
A;Accession: A23052
A;Molecule type: mRNA
A;Residues: 1-162 <HOP>
A;Cross-references: UNIPROT:P04926; UNIPARC:UPI000002BA56; GB:X01745; NID:g9859; PIDN:CA
C;Comment: This antigen may be secreted by the intraerythrocyte stage into the cytoplasm
tire protein from being secreted.
C;Comment: This antigen and the circumsporozoite protein appear to have a common epitope
C;Superfamily: plasmodium S-antigen
C;Keywords: sporozoite; surface antigen
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-162/Product: antigen 5.1 #status predicted <MAT>
F;76-101/Domain: transmembrane #status predicted <TM>

Query Match 5.1%; Score 6; DB 1; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
| | | | |
Db 91 VLLGGV 96

RESULT 181

ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A>Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A;Reference number: A89758; MUID:21311952; PMID:11418146
 A;Accession: E89955
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-164 <KUR>
 A;Cross-references: UNIPROT:Q99TFO; UNIPARC:UPI000013729E; GB:BA000018; PID:g13701508; E
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: SA1535
 C;Superfamily: thioredoxin peroxidase

Query Match 5.1%; Score 6; DB 2; Length 164;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
 Db 156 AALAA 161
 |||||

RESULT 186
 G83428
 hypothetical protein PA1746 [imported] - *Pseudomonas aeruginosa* (strain PAO1)
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C;Accession: G83428
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
 ; Lory, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic patho
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: G83428
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-168 <STO>
 A;Cross-references: UNIPROT:Q91221; UNIPARC:UPI00000C5477; GB:AE004600; GB:AE004091; NID
 A;Experimental source: strain PAO1
 C;Genetics:
 A;Gene: PA1746

Query Match 5.1%; Score 6; DB 2; Length 168;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 AQVIAH 80
 Db 14 AQVIAH 19
 |||||

RESULT 187
 F70660
 hypothetical protein Rv254c - *Mycobacterium tuberculosis* (strain H37RV)
 C;Species: *Mycobacterium tuberculosis*
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: F70660
 R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A>Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A;Reference number: A70500; MUID:98295987; PMID:9634230
 A;Accession: F70660
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-170 <COL>
 A;Cross-references: UNIPROT:P94999; UNIPARC:UPI000013BD34; GB:Z83863; GB:AL123456; NID:9
 A;Experimental source: strain H37RV

C;Genetics:
 A;Gene: Rv2554c
 C;Superfamily: Haemophilus influenzae conserved hypothetical protein HI0305

Query Match 5.1%; Score 6; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 Db 64 LAALAA 69
 |||||

RESULT 188
 F43256
 hypothetical protein 3 - *Thiobacillus ferrooxidans* plasmid pTF-PC2
 C;Species: *Thiobacillus ferrooxidans*
 C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C;Accession: F43256; S27626
 R;Rohrer, J.; Rawlings, D.E.
 J. Bacteriol. 174, 6230-6237, 1992
 A>Title: Sequence analysis and characterization of the mobilization region of a broad-ho
 A;Reference number: A43256; MUID:93015664; PMID:1400173
 A;Accession: F43256
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-170 <ROH>
 A;Cross-references: UNIPROT:P22902; UNIPARC:UPI000013B998; EMBL:M57717; NID:g154659; PID
 A;Note: sequence extracted from NCBI backbone (NCBIP:115315)
 C;Superfamily: *Thiobacillus ferrooxidans* plasmid pTF-PC2 hypothetical protein 3

Query Match 5.1%; Score 6; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 Db 105 LAALAA 110
 |||||

RESULT 189
 F82437
 probable rRNA methylase VCA0627 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1
 C;Species: *Vibrio cholerae*
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C;Accession: F82437
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: F82437
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-170 <HEI>
 A;Cross-references: UNIPROT:Q9KLM4; UNIPARC:UPI00000C35BA; GB:AE004392; GB:AE003853; NID
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VCA0627
 A;Map position: 2

Query Match 5.1%; Score 6; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 QAQVIA 79
 Db 153 QAQVIA 158
 |||||

RESULT 190
 AE1817

hypothetical protein all0085 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AE1817
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Kanehisa, K.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AE1817
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-170 <KUR>
 A:Cross-references: UNIPROT:Q8ZOK6; UNIPARC:UPI00000CDDC17; GB:BA000019; PIDN:BAE77609.1;
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all0085

Query Match 5.1%; Score 6; DB 2; Length 170;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGILQ 91
 |||||
 DB 60 VLGILQ 65

RESULT 191
 B75013
 phosphoribosylaminoimidazole carboxylase (pure) PAB1077 - Pyrococcus abyssi (strain Orsa)
 C:Species: Pyrococcus abyssi
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C:Accession: B75013
 R:anonymous, Genoscope
 submitted to the EMBL Data Library, July 1999
 A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
 A:Reference number: A75001
 A:Accession: B75013
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-174 <KAW>
 A:Cross-references: UNIPROT:Q9UV68; UNIPARC:UPI0000034562; GB:AJ248288; NID:11258796
 A:Experimental source: strain Orsa
 C:Genetics:
 A:Gene: purB; PAB1077
 C:Superfamily: phosphoribosylaminoimidazole carboxylase catalytic chain; phosphoribosylaminoimidazole carboxylase catalytic chain homology <PCC>

Query Match 5.1%; Score 6; DB 2; Length 174;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
 |||||
 DB 77 GVLAAL 82

RESULT 192
 B90988
 partial probable sensor kinase ECs2874 [similarity] - Escherichia coli (strain O157:H7,
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: B90988
 R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyana, K.; Han, C.G.; Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic map
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: B90988
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-177 <HAY>

Query Match

5.1%; Score 6; DB 2; Length 177;

A:Cross-references: UNIPROT:Q8X7L2; UNIPARC:UPI00000D0A3B; GB:BA000007; PIDN:BAE36297.1;
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs2874

Query Match 5.1%; Score 6; DB 2; Length 177;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||
 DB 137 LLGGVL 142

RESULT 193
 E85833
 partial probable sensor kinase Z3235 [imported] - Escherichia coli (strain O157:H7, substrain O157:H7)
 C:Species: Escherichia coli
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: E85833
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dialanta, E.; Potamouisis, K.; Apodaca, N.; Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: E85833
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-177 <STO>
 A:Cross-references: UNIPROT:Q8X7L2; UNIPARC:UPI00000D0A3B; GB:AB005174; NID:gl2516269; F:11258796
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z3235

Query Match 5.1%; Score 6; DB 2; Length 177;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||
 DB 137 LLGGVL 142

RESULT 194
 C95922
 hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSymb
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C:Accession: C95922
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan, Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymb magaplasmid from the N2-fixing endosymbiont
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: C95922
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-177 <KUR>

A:Cross-references: UNIPROT:Q92VQ6; UNIPARC:UPI00000CBSE3; GB:AL591985; PIDN:CAC49043.1;
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, M.; Neubalt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Smb21064
 A:Genome: plasmid

Query Match

5.1%; Score 6; DB 2; Length 177;

```
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ADLEVT 10
Db 36 ADLEVT 41

RESULT 195
H72697
hypothetical protein APE1000 - Aeropyrum pernix (strain KI)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: H72697
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; M
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: H72697
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-179 <KAW>
A:Cross-references: UNIPROT:Q9YDB3; UNIPARC:UPI000005DD8D; DDBJ:AP000060; NID:G5104188;
A:Experimental source: strain KI
C:Genetics:
C:Superfamily: Aeropyrum pernix hypothetical protein APE1000
A:Gene: APE1000

Query Match 5.1%; Score 6; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 40 LAALAA 45

RESULT 196
T50598
probable membrane protein [imported] - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50598
R;Redenbach, M.; Kiese, H.M.; Denapaite, D.; Eichner, A.; Cullum, J.; Kinashi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: T50598
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-180 <RED>
A:Cross-references: UNIPROT:Q9RKN9; UNIPARC:UPI00000DB660; EMBL:AL133220; PIDN:CAB61735.
A:Experimental source: strain A3(2)
C:Genetics:
A:Note: SCC75A.35

Query Match 5.1%; Score 6; DB 2; Length 180;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 19 GGVLA 24

RESULT 197
C83435
GTP cyclohydrolase I precursor PA1674 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: C83435
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
```

```
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-181 <STO>
A:Cross-references: UNIPROT:Q9I351; UNIPARC:UPI000012B278; GB:AE004595; GB:AE004091; NID
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: folE2; PA1674
C:Superfamily: GTP cyclohydrolase I

Query Match 5.1%; Score 6; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
Db 92 GKVLGL 97

RESULT 198
A86751
Prophage pi2 protein 02 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86751
R;Bolotin, A.; Wincker, P.; Manger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: A86751
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-181 <STO>
A:Cross-references: UNIPROT:Q9CGT3; UNIPARC:UPI000009BB32; GB:AE005176; PID:GI2723957; P
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: pi202

Query Match 5.1%; Score 6; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 APYIEQ 74
Db 15 APYIEQ 20

RESULT 199
F83175
conserved hypothetical protein PA3767 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83175
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83175
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-182 <STO>
A:Cross-references: UNIPROT:Q9HXM8; UNIPARC:UPI00000CSB00; GB:AE004795; GB:AE004091; NID
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA3767
```

Query Match 5.1%; Score 6; DB 2; Length 182;
 Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EBCSQ 68
 |||||
 Db 159 EBCSQ 164
 |||||

RESULT 200
 conserved hypothetical protein PA4612 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: D83069
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83069
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-183 <STO>
 A:Cross-references: UNIPROT:Q9HVH8; UNIPARC:UPI00000C5D8A; GB:AE004875; GB:AE004091; NID:11463916
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA4612

Query Match 5.1%; Score 6; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 14 LAALAA 19
 |||||

RESULT 201
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: F75255
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: F75255
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-185 <WHI>
 A:Cross-references: UNIPROT:Q9RRAL; UNIPARC:UPI00000C1B4F; GB:AE002089; GB:AE000513; NID:11463916
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2593
 A:Map position: 1
 C:Superfamily: Deinococcus radiodurans hypothetical protein DR1748

Query Match 5.1%; Score 6; DB 2; Length 185;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 157 LAALAA 162
 |||||

RESULT 202
 conserved hypothetical protein SP0719 [imported] - Streptococcus pneumoniae (strain TIGR C95083)
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C:Accession: C95083
 R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidison, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, N.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95083

hypothetical protein slr0808 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 C:Accession: S75549
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasud DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
 A:Reference number: S74322; MUID:97061201; PMID:8905231
 A:Accession: S75549
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-185 <KAN>
 A:Cross-references: UNIPROT:P74035; UNIPARC:UPI000013391E; EMBL:D90911; GB:AB001339; NID:11463916
 A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996
 C:Superfamily: 16S rRNA processing protein PA3744

Query Match 5.1%; Score 6; DB 2; Length 185;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVL 55
 |||||
 Db 152 PDKEVL 157
 |||||

RESULT 203
 GRP cyclohydrolase I precursor PA3438 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: D83217
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83217
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-186 <STO>
 A:Cross-references: UNIPROT:Q9HYG8; UNIPARC:UPI000012B276; GB:AE004764; GB:AE004091; NID:11463916
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: folE1; PA3438
 C:Superfamily: GTP cyclohydrolase I

Query Match 5.1%; Score 6; DB 2; Length 186;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
 |||||
 Db 94 GKVLGL 99
 |||||

RESULT 204
 conserved hypothetical protein SP0719 [imported] - Streptococcus pneumoniae (strain TIGR C95083)
 C:Species: Streptococcus pneumoniae
 C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C:Accession: C95083
 R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidison, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, N.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A:Reference number: A95000; MUID:21357209; PMID:11463916
 A:Accession: C95083

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-186 <KUR>
A;Cross-references: UNIPROT:Q97RS4; UNIPARC:UPI0000051561; GB:AE005672; PIDN:AAK74860.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0719

Query Match 5.1%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 77 VLAALA 82

RESULT 205

S49192

GCR 1 protein - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C;Accession: S49192

R;Parcment, C.; Hughes, D.M.; Lloyd, P.; Flavell, A.J.

submitted to the EMBL Data Library, May 1993

A;Description: A variety of different glycine repeats in *Drosophila* genes.

A;Reference number: S49192

A;Accession: S49192

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-188 <PAR>

A;Cross-references: UNIPROT:Q23709; UNIPARC:UPI0000084078; EMBL:X71976; NID:G2511633; PI

C;Genetics:

A;Gene: FlyBase:anon-Pen100DB

A;Cross-references: FlyBase:FBgn0003065

A;Introns: 6/3

C;Superfamily: glycine-rich cell wall structural protein 1

Query Match 5.1%; Score 6; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 12 LAALAA 17

RESULT 206

D82151

hypothetical protein VC1832 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)

C;Species: *Vibrio cholerae*

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: D82151

R;Heideberg, J.F.; Eiseen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82151

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-190 <HEI>

A;Cross-references: UNIPROT:Q9K15; UNIPARC:UPI00000C310F; GB:AE004259; GB:AE003852; NID

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC1832

A;Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
| | | | |
Db 133 LLGGVL 138

RESULT 207

AG2668

biotin synthesis BioY protein [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont

C;Species: *Agrobacterium tumefaciens*

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AG2668

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Monks, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AG2668

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-190 <KUR>

A;Cross-references: UNIPROT:Q8UHD3; UNIPARC:UPI00000D18F7; GB:AE008688; PIDN:AAI41765.1;

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: bioY

A;Map position: circular chromosome

Query Match 5.1%; Score 6; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 170 VLAALA 175

RESULT 208

E97450

probable biotin synthase [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)

C;Species: *Agrobacterium tumefaciens*

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: E97450

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: E97450

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-190 <KUR>

A;Cross-references: UNIPROT:Q8UHD3; UNIPARC:UPI00000D18F7; GB:AE007869; PIDN:AAK86558.1;

C;Genetics:

A;Gene: AGR_C_1360

A;Map position: circular chromosome

Query Match 5.1%; Score 6; DB 2; Length 190;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 170 VLAALA 175

RESULT 209

E87519

ThiJ/PfpI family protein [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: E87519

R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: E87519
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-191 <STO>
A:Cross-references: UNIPROT:Q9A6B2; UNIPARC:UPI00000C7631; GB:AE005673; NID:gl3423681; E
C:Genetics:
A:Gene: CC2182

Query Match 5.1%; Score 6; DB 2; Length 191;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
Db 46 GGVLA 51

RESULT 210

RnfB-related protein VC1016 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C:Species: *Vibrio cholerae*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82252
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragol, I.; Sellers, F. I., R.N.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: F82252
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-195 <HEI>
A:Cross-references: UNIPROT:Q9KT87; UNIPARC:UPI000013438B; GB:AE004183; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1016
A:Map position: 1
A:Superfamily: conserved hypothetical protein H11684; ferredoxin 2[4Fe-4S] homology

Query Match 5.1%; Score 6; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 11 LAALAA 16

RESULT 211

AB3092
probable phosphoheptose isomerase PA4425 [imported] - *Pseudomonas aeruginosa* (strain PAC
C:Species: *Pseudomonas aeruginosa*
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: AB3092
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Broman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic patho
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: AB3092
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <STO>
A:Cross-references: UNIPROT:Q9HVZ0; UNIPARC:UPI000012E842; GB:AE004857; NID
A:Experimental source: strain PAO1

C:Genetics:
A:Gene: PA4425
C:Superfamily: phosphoheptose isomerase

Query Match 5.1%; Score 6; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 PYIEQA 75
Db 27 PYIEQA 32

RESULT 212

A30013
plasma retinol-binding protein precursor - African clawed frog
C:Species: *Xenopus laevis* (African clawed frog)
C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 09-Jul-2004
C:Accession: A30013
R:McKearin, D.M.; Barton, M.C.; Keller, M.J.; Shapiro, D.J.
J. Biol. Chem. 262, 4939-4942, 1987
A:Title: Estrogen induces transcription of the *Xenopus laevis* serum retinol-binding pro
A:Reference number: A30013; MUID:87165916; PMID:3558378
A:Accession: A30013
A:Molecule type: mRNA
A:Residues: 1-197 <MCK>
A:Cross-references: UNIPROT:P06172; UNIPARC:UPI00001336E0; GB:J02718; NID:g214794; PIDN
C:Superfamily: lipocalin; lipocalin; lipocalin homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:35-194/Domain: lipocalin homology <LIP>
F:24-180,90-194,140-149/Disulfide bonds: #status predicted

Query Match 5.1%; Score 6; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
Db 4 KVLGLL 9

RESULT 213

F75436
conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: F75436
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <WHI>
A:Cross-references: UNIPROT:Q9RV00; UNIPARC:UPI000000C1892; GB:AE001960; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1099
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 15 GVLAAL 20

RESULT 214

G97950
conserved hypothetical protein spr0631 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: G97950
R:Hoekings, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczyk, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: G97950
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-197 <STO>
A:Cross-references: UNIPROT:Q8DQK3; UNIPARC:UPI00000834B4; GB:AE007317; PIDN:AAK99435.1;
C:Genetics:
A:Gene: spr0631

Query Match 5.1%; Score 6; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALA 25
| | | | |
Db 88 VLALA 93

RESULT 215
D87484
Glutamine amidotransferase, class I [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
C:Accession: D87484
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87484
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-198 <STO>
A:Cross-references: UNIPROT:Q9A729; UNIPARC:UPI00000C7542; GB:AE005673; NID:g13423346; F
C:Genetics:
A:Gene: CC1897
C:Superfamily: GMP synthase/anthranilate synthase; trpG homology

Query Match 5.1%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 ADLEVT 10
| | | | |
Db 139 ADLEVT 144

RESULT 216
C84237
hypothetical protein Vng0800h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84237
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: C84237

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-198 <STO>
A:Cross-references: UNIPROT:Q9HR95; UNIPARC:UPI0000063725; GB:AE004437; NID:g10580373; P
C:Genetics:
A:Gene: VNG0800H

Query Match 5.1%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
| | | | |
Db 72 GGVLA 77

RESULT 217
AE3435
hypothetical protein BME11467 [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AE3435
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Golcsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3435
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-198 <STO>
A:Cross-references: UNIPROT:Q8YFQ3; UNIPROT:Q8G265; UNIPARC:UPI00000580A1; GB:AE008917;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11467
A:Map position: I

Query Match 5.1%; Score 6; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 5 LAALAA 10

RESULT 218
AB0236
probable Na(+)-translocating NADH-quinone reductase chain E (EC 1.6.5.-) [imported] - Ye
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 31-Dec-2004
C:Accession: AB0236
R:Packhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0236
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-199 <STO>
A:Cross-references: UNIPROT:Q9ZC45; UNIPARC:UPI0000130461; GB:AL590842; PIDN:CAC90750.1;
C:Genetics:
A:Gene: YP01935
C:Superfamily: Na(+)-translocating NADH-quinone reductase, subunit D/E
C:Keywords: oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24

Db 86 GVLAAL 91
|||||

RESULT 219

T47023
hypothetical protein [imported] - Yersinia pestis
C:Species: Yersinia pestis
C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 31-Dec-2004
C:Accession: T47023
R:Buchrieser, C.; Rusniok, C.; Couve, E.; Frangeul, L.; Billault, A.; Kunst, F.; Carniel
submitted to the EMBL Data Library, October 1998
A:Description: DNA sequence of the 102 kbases unstable region of Yersinia pestis.
A:Reference number: Z4348
A:Accession: T47023
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-199 <BUC>
A:Cross-references: UNIPROT:Q9ZC45; UNIPARC:UPI0000130461; EMBL:AL031866; PIDN:CAA21366.
A:Experimental source: strain 6/69
C:Superfamily: Na(+)-translocating NADH-quinone reductase, subunit D/E

Query Match 5.1%; Score 6; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||

Db 86 GVLAAL 91

RESULT 220

S15164
hypothetical protein 2 - Pseudomonas syringae pv. phaseolicola insertion sequence IS801
C:Species: Pseudomonas syringae pv. phaseolicola
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: S15164; S18262
R:Romantschuk, M.; Richter, G.Y.; Mukhopadhyay, P.; Mills, D.
Mol. Microbiol. 5, 617-622, 1991
A>Title: IS801, an insertion sequence element isolated from Pseudomonas syringae pathovar
A:Reference number: S15163; MUID:91260445; PMID:1646375
A:Accession: S15164
A:Molecule type: DNA
A:Residues: 1-199 <ROM>
A:Cross-references: UNIPROT:Q99178; UNIPARC:UPI0000179224; EMBL:X57269
A:Experimental source: strain LR781
R:Mills, D.
submitted to the EMBL Data Library, January 1991

A:Reference number: S18262
A:Accession: S18262
A:Molecule type: DNA
A:Residues: 1-33, 'p', '35-199 <MIL>
A:Cross-references: UNIPARC:UPI0000084E37; EMBL:X57269; NID:945830; PIDN:CAA40541.1; PID

A:Experimental source: strain LR781
C:Genetics:
A:Mobile element: insertion sequence IS801
C:Superfamily: Pseudomonas syringae insertion sequence IS801 hypothetical protein 2

Query Match 5.1%; Score 6; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 76 QVIAHQ 81
|||||

Db 111 QVIAHQ 116

RESULT 221

T36622
hypothetical protein SCH35_37c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36622

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21610
A:Accession: T36622
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-199 <OLI>
A:Cross-references: UNIPROT:Q9X8W9; UNIPARC:UPI00000DB0C7; EMBL:AL078610; PIDN:CAB44390
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH35.37c

Query Match 5.1%; Score 6; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

Db 87 LAALAA 92

RESULT 222

T43151
hypothetical protein - fission yeast (Schizosaccharomyces pombe) (fragment)
C:Species: Schizosaccharomyces pombe
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43151
R:Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A>Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A:Reference number: Z17323; MUID:98162722; PMID:9501991
A:Accession: T43151
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-201 <YOS>
A:Cross-references: UNIPROT:P78893; UNIPARC:UPI000006A723; EMBL:D89244; NID:gl749695; P

A:Experimental source: strain PR745
Query Match 5.1%; Score 6; DB 2; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

Db 114 LAALAA 119

RESULT 223

F87540
hypothetical protein CC2351 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: F87540
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: F87540
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <STO>
A:Cross-references: UNIPROT:Q9ASU6; UNIPARC:UPI00000C76CB; GB:AE005673; NID:gl3423878; P

Query Match 5.1%; Score 6; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
|||||

Db 129 GGVLA 134

RESULT 224
A99865
hypothetical protein ECs1889 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: A99865
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. Gasegawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference number: A99629; PMID:11258796
A:Accession: A99865
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <HAY>
A:Cross-references: UNIPROT:Q8X8I9; UNIPARC:UPI0000165440; GB:BA0000007; PIDN:BAB35312.1
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs1889

Query Match 5.1%; Score 6; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 VTTSTW 14
|||||

Db 120 VTTSTW 125

RESULT 225
H85753
partial probable periplasmic transport protein 22473 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H85753
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.W.; Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamou, K.; Apodaca, Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; PMID:21074935; PMID:11206551
A:Accession: H85753
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-204 <STO>
A:Cross-references: UNIPROT:Q8X8I9; UNIPARC:UPI00000D0923; GB:AE005174; NID:G12515459; H85753
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: 22473

Query Match 5.1%; Score 6; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 VTTSTW 14
|||||

Db 121 VTTSTW 126

RESULT 226
AF2299
cobalt transport ATP-binding protein cbiO [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AF2299
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; PMID:21595285; PMID:11759840

Query Match 5.1%; Score 6; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A:Accession: AF2299
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <KUR>
A:Cross-references: UNIPROT:Q8YQ85; UNIPARC:UPI00000CE973; GB:BA0000019; PIDN:BAB75648.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: cbiO
C:Superfamily: short-chain ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 5.1%; Score 6; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LALAA 26
|||||

Db 113 LALAA 118

RESULT 227
AC1997
hypothetical protein alr1529 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC1997
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; PMID:21595285; PMID:11759840
A:Accession: AC1997
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-206 <KUR>
A:Cross-references: UNIPROT:Q8YWS4; UNIPARC:UPI00000CE106; GB:BA0000019; PIDN:BAB77895.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1529

Query Match 5.1%; Score 6; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 69 APYIEQ 74
|||||

Db 126 APYIEQ 131

RESULT 228
D81999
glucose inhibited division protein B homolog NMA0077 [imported] - Neisseria meningitidis
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: D81999
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel, J.; Holroyd, S.; Jogle, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream, Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: AB1775; PMID:20222556; PMID:10761919
A:Accession: D81999
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-207 <PAR>
A:Cross-references: UNIPROT:Q9JX38; UNIPARC:UPI000012B47A; GB:AL162752; GB:AL157959; NID:Z2491
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: gldB; NMA0077
C:Superfamily: gldB protein

Query Match 5.1%; Score 6; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QOAVIE 101
 |||||
 Db 109 QOAVIE 114

RESULT 229

C81227
 Glucose inhibited division protein B NMB0190 [imported] - Neisseria meningitidis (strain
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: C81227
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: C81227
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-207 <TET>
 A:Cross-references: UNIPROT:Q9KLQ3; UNIPARC:UPI000012B47B; GB:AE002376; GB:AE002098; NID:
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB0190
 C:Superfamily: gidB protein

Query Match 5.1%; Score 6; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QOAVIE 101
 |||||
 Db 109 QOAVIE 114

RESULT 230

B69797
 conserved hypothetical protein yeev - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: B69797
 R:Kumat, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
 C.; Bron, S.; Broutillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Gall
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Maueel
 Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Togmon, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: B69797
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-208 <KUN>
 A:Cross-references: UNIPROT:O31525; UNIPARC:UPI0000060059; GB:Z99107; GB:AL009126; NID:9
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yeev
 C:Superfamily: Bacillus subtilis conserved hypothetical protein yeev

Query Match 5.1%; Score 6; DB 1; Length 208;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVL 21
 |||||
 Db 31 LGGVL 36

RESULT 231

C96773
 hypothetical protein FIM20.12 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: C96773
 R:Rheologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.P.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luo, J.S.; Maity, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: C96773
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-208 <STO>
 A:Cross-references: UNIPROT:Q9CA70; UNIPARC:UPI000009E0DA; GB:AE005173; NID:96539246; P
 C:Genetics:
 A:Gene: FIM20.12
 A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 208;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAL 24
 |||||
 Db 93 GVLAL 98

RESULT 232

B70738
 hypothetical protein Rv3421c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: B70738
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: B70738
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-211 <COL>
 A:Cross-references: UNIPROT:Q50707; UNIPARC:UPI000013C27B; GB:Z77165; GB:AL123456; NID:9
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv3421c
 C:Superfamily: inactive homolog of metal-dependent protease

Query Match 5.1%; Score 6; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLALA 25
 |||||
 Db 51 VLALA 56

```
RESULT 233
AG0854
Probable sugar aldolase [imported] - Salmonella enterica subsp. enterica serovar Typhi
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: This species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 05-Oct-2004
C;Accession: AG0854
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0854
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-212 <PAR>
A;Cross-references: UNIPARC:UPI000005A388; GB:AL513382; PIDN:CAD06022.1; PID:gl6503991;
C;Genetics:
A;Gene: STV3041
C;Superfamily: Class II aldolase/adducin, N-terminal

Query Match          5.1%; Score 6; DB 2; Length 212;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
    |||||
Db 150 LAALAA 155

RESULT 234
B69493
Hypothetical protein AF1947 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: B69493
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Corton, M.D.; Spriggs, T.; Attiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: B69493
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-212 <KLE>
A;Cross-references: UNIPROT:Q28332; UNIPARC:UPI0000056B2B; GB:AE0000969; GB:AE000782; NID

Query Match          5.1%; Score 6; DB 2; Length 212;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
    |||||
Db 197 VLAALA 202

RESULT 235
S27449
Sc14 protein - bracket fungus (Schizophyllum commune)
C;Species: Schizophyllum commune
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004
C;Accession: S27449
R;Schuren, F.H.; Kothe, E.M.; Wessels, J.G.
submitted to the EMBL Data Library, April 1992
A;Reference number: S27449
A;Accession: S27449
A;Status: preliminary
A;Molecule type: mRNA
```

```
A;Residues: 1-214 <SCH>
A;Cross-references: UNIPROT:P35795; UNIPARC:UPI00000135B9; EMBL:M81723; NID:gl69866; PID
C;Superfamily: pathogenesis-related leaf protein

Query Match          5.1%; Score 6; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
    |||||
Db 9 LAALAA 14

RESULT 236
AB3163
Probable carboxylesterase PA3859 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: AB3163
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: AB3163
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-215 <STO>
A;Cross-references: UNIPROT:Q9HXE7; UNIPARC:UPI0000005B47; GB:AE004803; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA3859

Query Match          5.1%; Score 6; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
    |||||
Db 131 LGGVLA 136

RESULT 237
B83604
Hypothetical protein PA0335 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2004
C;Accession: B83604
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: B83604
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-217 <STO>
A;Cross-references: UNIPROT:Q916F6; UNIPARC:UPI000000C5012; GB:AE004471; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0335
C;Superfamily: phosphoserine phosphatase

Query Match          5.1%; Score 6; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 99 VIEPIV 104
    |||||
Db 84 VIEPIV 89
```

RESULT 238

T49885
peptide methionine sulfoxide reductase-like protein - Arabidopsis thaliana
N:Alternate names: protein T211.170
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49885
R:Sevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24493
A:Accession: T49885
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <BEV>
A:Cross-references: UNIPROT:Q9LY15; UNIPARC:UPI00000AA3BE; EMBL:AL163912; GSPDB:GN000063;
A:Experimental source: cultivar Columbia; BAC clone T211
C:Genetics:
A:Gene: ATSP-T211.170
A:Map position: 5
A:introns: 135/3
C:Superfamily: peptide methionine sulfoxide reductase

Query Match 5.1%; Score 6; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 KP41VP 50
|||||
Db 30 KP41VP 35

RESULT 239

B84335
hypothetical protein Vng1843c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84335
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: B84160; MUID:20504483; PMID:11016950
A:Accession: B84335
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-218 <STO>
A:Cross-references: UNIPROT:Q9HP21; UNIPARC:UPI00000639A7; GB:AE004437; NID:g10581289; H
C:Genetics:
A:Gene: VNG1843C

Query Match 5.1%; Score 6; DB 2; Length 218;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 168 VLAALA 173

RESULT 240

CB3454
probable transcription regulator PA1526 [imported] - Pseudomonas aeruginosa (strain PA01
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: CB3454
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83454

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-219 <STO>

A:Cross-references: UNIPROT:Q91317; UNIPARC:UPI00000C53C4; GB:AE004581; GB:AE004091; NII

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA1526

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 219;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

|||||

Db 212 VLAALA 217

RESULT 241

C95009
potassium uptake protein, Trk family [imported] - Streptococcus pneumoniae (strain TIGR4
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C:Accession: C95009
R:Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, B.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95009
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-221 <KUR>
A:Cross-references: UNIPROT:Q97T72; UNIPARC:UPI0000051B8; GB:AE005672; PIDN:AAK74268.1;
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0079
C:Superfamily: conserved hypothetical protein MG323

Query Match 5.1%; Score 6; DB 2; Length 221;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

|||||

Db 18 VLAALA 23

RESULT 242

AH2116
hypothetical protein all2487 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2116
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2116
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-224 <KUR>
A:Cross-references: UNIPROT:Q8YU67; UNIPARC:UPI00000CE45D; GB:BA000019; PIDN:BAH74186.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all2487

Query Match 5.1%; Score 6; DB 2; Length 224;

Best Local Similarity 100.0%; Pred. No. 3.6e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 22 AALAA 27
Db 91 AALAA 96
|||||

RESULT 243
AH0683
hypothetical protein Atu0751 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AJ2668
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.
erag, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AJ2668
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <KUR>
A:Cross-references: UNIPROT:Q8UHD1; UNIPARC:UPI000016451B; GB:AE008688; PIDN:AAL41767.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0751
A:Map position: circular chromosome
C:Superfamily: syphilis spirochete probable ABC transporter nraA; ATP-binding cassette h

Query Match 5.1%; Score 6; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 141 LAALAA 146
|||||

RESULT 244
JC5327
adhesin complex 25K protein precursor - Eikenella corrodens
N:Alternate names: LecA protein
C:Species: Eikenella corrodens
C>Date: 15-May-1997 #sequence_revision 12-Sep-1997 #text_change 09-Jul-2004
C:Accession: JC5327; PC4312
R:Azakami, H.; Yumoto, H.; Nakae, H.; Matsuo, T.; Ebisu, S.
Gene 180, 207-212, 1996
A:Title: Molecular analysis of the gene encoding a protein component of the Eikenella co
A:Reference number: JC5327; MUID:97128828; PMID:8973368
A:Accession: JC5327
A:Molecule type: DNA
A:Residues: 1-226 <AZA1>
A:Cross-references: UNIPROT:P94741; UNIPARC:UPI000017CBE6; DDBJ:D78153; NID:gi1817527; PI
A:Accession: PC4312
A:Molecule type: protein
A:Residues: 23-32 <AZA2>
A:Cross-references: UNIPARC:UPI000017CBE7
C:Comment: This protein has 7 membrane-spanning regions and other characteristics in com
C:Genetics:
A:Gene: lecA
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-226/Product: adhesin complex 25K protein #status predicted <MAT>

Query Match 5.1%; Score 6; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 91 AALAA 96
|||||

Db 8 VLAALA 13
|||||

RESULT 245
AH0683
hypothetical protein STV1595 [imported] - Salmonella enterica subsp. enterica serovar Typh
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH0683
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH0683
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-227 <PAR>
A:Cross-references: UNIPARC:UPI000005956B; GB:AL513382; PIDN:CAD01841.1; PID:gl6502685;
C:Genetics:
A:Gene: STV1595

Query Match 5.1%; Score 6; DB 2; Length 227;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
Db 105 AALAA 110
|||||

RESULT 246
G97880
hypothetical protein trkA [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: G97880
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Bargett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Mateushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: G97880
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-228 <KUR>
A:Cross-references: UNIPROT:Q8DRK4; UNIPARC:UPI00000E33A2; GB:AE007317; PIDN:AAK98875.1;
C:Genetics:
A:Gene: trkA
C:Superfamily: conserved hypothetical protein MG323

Query Match 5.1%; Score 6; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 25 VLAALA 30
|||||

RESULT 247
DB3441
two-component response regulator KdpE PA1637 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: DB3441
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, K.R.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83441
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <STO>
A:Cross-references: UNIPROT:Q9I388; UNIPARC:UPI00000C541D; GB:AE004591; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: kdpE; PA1637
C:Superfamily: ompR protein; response regulator homology

Query Match 5.1%; Score 6; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
Db 163 VLAALA 168
|||||

RESULT 248
T36672
membrane-spanning protein - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36672
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A:Reference number: Z21611
A:Accession: T36672
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-230 <SEQ>
A:Cross-references: UNIPROT:Q9X916; UNIPARC:UPI00000DAP0C; EMBL:AL035636; PMID:CA838488.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCH5.14C

Query Match 5.1%; Score 6; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 131 LAALAA 136
|||||

RESULT 249
DB4006
lytic transglycosylase BH2852 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: DB4006
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: DB4006
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <STO>
A:Cross-references: UNIPROT:Q9K900; UNIPARC:UPI00000C4016; GB:AP001516; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2852

Query Match 5.1%; Score 6; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 QAQVIA 79
Db 74 QAQVIA 79
|||||

RESULT 250

T35302
hypothetical protein SC5F7.15 - *Streptomyces coelicolor*
C:Species: *Streptomyces coelicolor*
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35302
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21574
A:Accession: T35302
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-234 <SEQ>
A:Cross-references: UNIPROT:Q9S2Q0; UNIPARC:UPI00000DB29B; EMBL:AL096872; PMID:CA851270
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC5F7.15

Query Match 5.1%; Score 6; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 WVLGG 19
Db 125 WVLGG 130
|||||

RESULT 251

T06462
glutathione peroxidase (EC 1.11.1.9) precursor - garden pea
N:Alternate names: phospholipid glutathione peroxidase
C:Species: *Pisum sativum* (garden pea)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06462
R:Mullineaux, P.M.; Karpinski, S.; Jimenez, A.; Cleary, S.P.; Robinson, C.; Creissen, G.
Plant J. 13, 375-379, 1998
A:Title: Identification of cDNAs encoding plastid-targeted glutathione peroxidase.
A:Reference number: Z15696; MUID:98345965; PMID:9680987
A:Accession: T06462
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-236 <MUL>
A:Cross-references: UNIPROT:O24296; UNIPARC:UPI000012BC10; EMBL:AJ7000508; NID:G2632108;
A:Experimental source: cv. Birte
C:Genetics:
A:Map position: 1
A:Genome: nuclear
C:Superfamily: glutathione peroxidase
C:Keywords: chloroplast; oxidoreductase; selenocysteine

Query Match 5.1%; Score 6; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 FKGVKL 87
Db 97 FKGVKL 102
|||||

RESULT 252

T05973
permatin homolog PR5 - barley
C:Species: *Hordeum vulgare* (barley)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05973
R:Collinge, D.B.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z15479

A/Accession: T05973
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-237 <COL>
A/Cross-references: UNIPROT:Q23997; UNIPARC:UPI000009CB3D; EMBL:AJ001268; PIDN:CAA04642.
A/Experimental source: cv, Pallas
C/Superfamily: thaumatin I

Query Match 5.1%; Score 6; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 15 LAALAA 20

RESULT 253
B64643
Hypothetical protein HP0986 - Helicobacter pylori (strain 26695)
C/Species: Helicobacter pylori
C/Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C/Accession: B64643
R/Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
A/Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A/Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A/Reference number: A64520; MUID:97394467; PMID:9252185
A/Accession: B64643
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-237 <TOM>
A/Cross-references: UNIPROT:Q25636; UNIPARC:UPI00000C08D0; GB:AE000607; GB:AE000511; NID:97394467
C/Genetics:
A/Start codon: GTG

Query Match 5.1%; Score 6; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 176 GVLAAL 181

RESULT 254
F70606
Probable pknM protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: F70606
R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gencies, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: F70606
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-237 <COL>
A/Cross-references: UNIPROT:P96858; UNIPARC:UPI0000031889; GB:Z92774; GB:AL123456; NID:98295987
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: pknM

Query Match 5.1%; Score 6; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 5 LAALAA 10

RESULT 255
C42606
orfB 3', to orf405 - Saccharopolyspora erythraea (fragment)
C/Species: Saccharopolyspora erythraea
C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Dec-2004
C/Accession: C42606
R/Andersen, J.F.; Hutchinson, C.R.
J. Bacteriol. 174, 725-735, 1992
A/Title: Characterization of Saccharopolyspora erythraea cytochrome P-450 genes and enzyme
A/Reference number: A42606; MUID:92121109; PMID:1732208
A/Contents: NRRL2338
A/Accession: C42606
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-237 <AND>
A/Cross-references: UNIPROT:P34698; UNIPROT:Q33630; UNIPARC:UPI00000B9C45; GB:M83110; NCI:77485)
A/Note: sequence extracted from NCBI backbone (NCBIN:77481, NCBIN:77476, NCBIP:77485)

Query Match 5.1%; Score 6; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 71 GVLAAL 76

RESULT 256
AC2874
chitooligosaccharide deacetylase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 31-Dec-2004
C/Accession: AC2874
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Eraso, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AC2874
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-238 <KUR>
A/Cross-references: UNIPROT:Q8UCR2; UNIPARC:UPI00000D1ECA; GB:AE008688; PIDN:AAL43409.1;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu2421
A/Map position: circular chromosome
C/Superfamily: nodulation protein nodB

Query Match 5.1%; Score 6; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAAL 25
| | | | |
Db 12 VLAAAL 17

RESULT 257
D97650
conserved hypothetical protein BH1302 (AP001511) [imported] - Agrobacterium tumefaciens
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 31-Dec-2004
C/Accession: D97650
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, P.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: D97650
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-238 <KUR>
 A:Cross-references: UNIPROT:Q8UCR2; UNIPARC:UPI00000D1ECA; GB:AE007869; PIDN:AAK88157.1;
 C:Genetics:
 A:Gene: AGR_C 4393
 A:Map position: circular chromosome
 C:Superfamily: nodulation protein nodB

Query Match 5.1%; Score 6; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
 |||||
 Db 12 VLAALA 17

RESULT 258
 T31218
 hypothetical protein 766 - Sphingomonas aromaticivorans plasmid pML1
 C:Species: Sphingomonas aromaticivorans
 C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
 C:Accession: T31218
 R:Romine, M.P.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.; C. submitted to the EMBL Data Library, July 1998
 A:Description: Complete sequence of a 184 kb catabolic plasmid from Sphingomonas aromaticivorans
 A:Reference number: Z20992
 A:Accession: T31218
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-238 <ROM>
 A:Cross-references: UNIPROT:085927; UNIPARC:UPI000005C9B3; EMBL:AF079317; NID:G3378261;
 C:Genetics:
 A:Genome: plasmid pML1
 A:Note: orf766
 C:Superfamily: Sphingomonas aromaticivorans hypothetical protein 766

Query Match 5.1%; Score 6; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAA 27
 |||||
 Db 159 AALAA 164

RESULT 259
 C75379
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: C75379
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M. S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: C75379
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-238 <WHI>
 A:Cross-references: UNIPROT:Q9RU22; UNIPARC:UPI00000C1963; GB:AE002001; GB:AE000513; NID:G3378261;
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1574
 A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 238;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 29 LAALAA 34

RESULT 260

G82383
 oxidoreductase, short-chain dehydrogenase/reductase family VCA1057 [imported] - Vibrio cholerae
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 05-Oct-2004
 C:Accession: G82383
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: G82383
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-239 <HEI>
 A:Cross-references: UNIPROT:Q9KKP5; UNIPARC:UPI00000C3743; GB:AE004432; GB:AE003853; NID:G3378261;
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA1057
 A:Map position: 2
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

QY 18 GGVLA 23
 |||||
 Db 203 GGVLA 208

RESULT 261

JC7768
 blue fluorescent protein, BFPV - Vibrio vulnificus
 C:Species: Vibrio vulnificus
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 05-Oct-2004
 C:Accession: JC7768
 R:Su, J.H.; Chuang, Y.C.; Tsai, Y.C.; Chang, M.C.
 A>Title: Cloning and characterization of a blue fluorescent protein from Vibrio vulnificus
 A:Reference number: JC7768; PMID:11554735
 A:Accession: JC7768
 A:Molecule type: DNA
 A:Residues: 1-239 <SUA>
 A:Cross-references: UNIPROT:Q9F172; UNIPARC:UPI00000597E5; GB:AF080431
 A:Experimental source: strain CKM-1
 C:Comment: This protein that shows significant homology to the short-chain dehydrogenase irradiated by UV light source.
 C:Genetics:
 A:Gene: bfpv
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
 |||||
 Db 203 GGVLA 208

RESULT 262

JC7768
 blue fluorescent protein, BFPV - Vibrio vulnificus
 C:Species: Vibrio vulnificus
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 05-Oct-2004
 C:Accession: JC7768
 R:Su, J.H.; Chuang, Y.C.; Tsai, Y.C.; Chang, M.C.
 A>Title: Cloning and characterization of a blue fluorescent protein from Vibrio vulnificus
 A:Reference number: JC7768; PMID:11554735
 A:Accession: JC7768
 A:Molecule type: DNA
 A:Residues: 1-239 <SUA>
 A:Cross-references: UNIPROT:Q9F172; UNIPARC:UPI00000597E5; GB:AF080431
 A:Experimental source: strain CKM-1
 C:Comment: This protein that shows significant homology to the short-chain dehydrogenase irradiated by UV light source.
 C:Genetics:
 A:Gene: bfpv
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
 |||||
 Db 203 GGVLA 208

RESULT 262

JC7768
 blue fluorescent protein, BFPV - Vibrio vulnificus
 C:Species: Vibrio vulnificus
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 05-Oct-2004
 C:Accession: JC7768
 R:Su, J.H.; Chuang, Y.C.; Tsai, Y.C.; Chang, M.C.
 A>Title: Cloning and characterization of a blue fluorescent protein from Vibrio vulnificus
 A:Reference number: JC7768; PMID:11554735
 A:Accession: JC7768
 A:Molecule type: DNA
 A:Residues: 1-239 <SUA>
 A:Cross-references: UNIPROT:Q9F172; UNIPARC:UPI00000597E5; GB:AF080431
 A:Experimental source: strain CKM-1
 C:Comment: This protein that shows significant homology to the short-chain dehydrogenase irradiated by UV light source.
 C:Genetics:
 A:Gene: bfpv
 C:Superfamily: short-chain dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
 |||||
 Db 203 GGVLA 208

RESULT 262

T34950
hypothetical protein SC4A10.13c - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 31-Dec-2004
C/Accession: T34950
R/Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21563
A/Accession: T34950
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-239 <SAU>
A/Cross-references: UNIPROT:Q9S2X1; UNIPARC:UPI00000DB2D0; EMBL:AL109663; PIDN:CAB51989.
A/Experimental source: strain A3(2)
C/Genetics:
C/Superfamily: Class III pyridoxal 5'-phosphate dependent enzyme, YBL036c type

Query Match 5.1%; Score 6; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MSADLE 8
|||||
Db 210 MSADLE 215

RESULT 263
F86738
hypothetical protein yjdf [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86738
R/Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissensbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: F86738
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-239 <STO>
A/Cross-references: UNIPROT:Q9CH31; UNIPARC:UPI00000D4411; GB:AE005176; PID:G12723843; P
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: yjdf

Query Match 5.1%; Score 6; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 EVTTST 13
|||||
Db 57 EVTTST 62

RESULT 264
A41797
Ig light chain - sandbar shark
C/Species: Carcharhinus plumbeus (sandbar shark)
C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C/Accession: A41797
R/Hohman, V.S.; Schluter, S.P.; Marchalonis, J.J.
Proc. Natl. Acad. Sci. U.S.A. 89, 276-280, 1992
A/Title: Complete sequence of a cDNA clone specifying sandbar shark immunoglobulin light
A/Reference number: A41797; MUID:92108036; PMID:1729697
A/Accession: A41797
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-240 <HOH>
A/Cross-references: UNIPARC:UPI00001152D0; GB:M81314; NID:G212937; PIDN:AAA50806.1; PID:
A/Experimental source: spleen
A/Note: sequence extracted from NCBI backbone (NCBIN:74783, NCBIP:74784)
C/Superfamily: immunoglobulin V region; immunoglobulin homology

T34950
hypothetical protein SC4A10.13c - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 31-Dec-2004
C/Accession: T34950
R/Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21563
A/Accession: T34950
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-239 <SAU>
A/Cross-references: UNIPROT:Q9S2X1; UNIPARC:UPI00000DB2D0; EMBL:AL109663; PIDN:CAB51989.
A/Experimental source: strain A3(2)
C/Genetics:
C/Superfamily: Class III pyridoxal 5'-phosphate dependent enzyme, YBL036c type

Query Match 5.1%; Score 6; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MSADLE 8
|||||
Db 210 MSADLE 215

RESULT 263
F86738
hypothetical protein yjdf [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C/Accession: F86738
R/Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissensbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A/Reference number: A86625; MUID:21235186; PMID:11337471
A/Accession: F86738
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-239 <STO>
A/Cross-references: UNIPROT:Q9CH31; UNIPARC:UPI00000D4411; GB:AE005176; PID:G12723843; P
A/Experimental source: strain IL1403
C/Genetics:
A/Gene: yjdf

Query Match 5.1%; Score 6; DB 2; Length 239;
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 EVTTST 13
|||||
Db 57 EVTTST 62

RESULT 264
A41797
Ig light chain - sandbar shark
C/Species: Carcharhinus plumbeus (sandbar shark)
C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C/Accession: A41797
R/Hohman, V.S.; Schluter, S.P.; Marchalonis, J.J.
Proc. Natl. Acad. Sci. U.S.A. 89, 276-280, 1992
A/Title: Complete sequence of a cDNA clone specifying sandbar shark immunoglobulin light
A/Reference number: A41797; MUID:92108036; PMID:1729697
A/Accession: A41797
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-240 <HOH>
A/Cross-references: UNIPARC:UPI00001152D0; GB:M81314; NID:G212937; PIDN:AAA50806.1; PID:
A/Experimental source: spleen
A/Note: sequence extracted from NCBI backbone (NCBIN:74783, NCBIP:74784)
C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: immunoglobulin

Query Match 5.1%; Score 6; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||||
Db 6 GVLAAL 11

RESULT 265

H84197
hypothetical protein Vng0391c [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: H84197
R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: H84197
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-240 <STO>
A/Cross-references: UNIPROT:Q9HS60; UNIPARC:UPI0000063604; GB:AE004437; NID:G10580006; P
C/Genetics:
A/Gene: VNG0391C
C/Superfamily: hypothetical protein M0570

Query Match 5.1%; Score 6; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 181 LAALAA 186

RESULT 266

T33165
hypothetical protein T26C12.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C/Accession: T33165
R/Gattung, S.; Scheet, P.; Harper, M.
submitted to the EMBL Data Library, May 1998
A/Description: The sequence of C. elegans cosmid T26C12.
A/Reference number: Z21296
A/Accession: T33165
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-240 <GAT>
A/Cross-references: UNIPROT:O61857; UNIPARC:UPI000007DA99; EMBL:AF067623; PIDN:AAC17554.
A/Experimental source: strain Bristol N2; clone T26C12
C/Genetics:
A/Gene: CESP:T26C12.2
A/Map position: 4
A/Introns: 120/3; 206/1

Query Match 5.1%; Score 6; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVL 21
|||||
Db 72 LLGGVL 77

RESULT 267

C81971
hypothetical protein NMA0528 [imported] - Neisseria meningitidis (strain Z2491 serogroup R)
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 31-Dec-2004
C:Accession: C81971
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: C81971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-241 <PAR>
A:Cross-references: UNIPROT:Q9JW64; UNIPARC:UPI000000C49BB; GB:AL162753; GB:AL157959; NID
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA0528
C:Superfamily: methyltransferase, slr0722 type

Query Match 5.1%; Score 6; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 227 LAALAA 232

RESULT 268
H81026
conserved hypothetical protein NMB1925 [imported] - Neisseria meningitidis (strain MC58)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Dec-2004
C:Accession: H81026
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: H81026
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-241 <TET>
A:Cross-references: UNIPROT:Q9JX08; UNIPARC:UPI000000C4815; GB:AE002541; GB:AE002098; NID
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1925
C:Superfamily: methyltransferase, slr0722 type

Query Match 5.1%; Score 6; DB 2; Length 241;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 227 LAALAA 232

RESULT 269
G75607
phosphoesterase-related protein DRA0124 [similarity] - Deinococcus radiodurans (strain R
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Sep-2003
C:Accession: G75607
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: G75607
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <WHI>
A:Cross-references: UNIPARC:UPI000000C163E; GB:AE001862; GB:AE001825; NID:G6460468; PIDN
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0124
A:Map position: 2
C:Superfamily: phosphoesterase, MJ0912 type; phosphoesterase core homology

Query Match 5.1%; Score 6; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VVIVGH 39
|||||
Db 153 VVIVGH 158

RESULT 270
F69553
enoyl-CoA hydratase (fad-5) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Aug-2004
C:Accession: F69553
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: F69553
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-243 <KLE>
A:Cross-references: UNIPROT:Q30242; UNIPARC:UPI0000057293; GB:AE001107; GB:AE000782; NID
C:Superfamily: Naphthoate synthase; enoyl-CoA hydratase homology
P:23-177/Domain: enoyl-CoA hydratase homology <ECH>

Query Match 5.1%; Score 6; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVGL 89
|||||
Db 171 GKVGL 176

RESULT 271
G75422
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: G75422
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <WHI>
A:Cross-references: UNIPROT:Q9RV09; UNIPARC:UPI000000C18CB; GB:AE001970; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRI221
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 68 LAALAA 73

RESULT 272
D64322
ribosomal protein L2 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 07-Jul-2003
C:Accession: D64322
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Raun, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: D64322
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-246 <BUL>
A:Cross-references: UNIPARC:UPI00001772DA; GB:U67474; GB:L77117; NID:gi590921; PID:gi592220
C:Genetics:
A:Map position: FOR180572-181312
A:Start codon: GTG
C:Superfamily: ribosomal protein L2/L8

Query Match 5.1%; Score 6; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLG 88
|||||
Db 43 KGKVLG 48

RESULT 273
AD0757
cobalamin (5'-phosphate) synthase (EC 2.-.-.-) [imported] - Salmonella enterica subsp. e
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 05-Oct-2004
C:Accession: AD0757
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-247 <PAR>
A:Cross-references: UNIPARC:UPI0000059C85; GB:AL513382; PIDN:CAD02378.1; PID:gi6503252;
C:Genetics:
A:Gene: STY2220
C:Superfamily: Cobalamin synthase
C:Keywords: transferase

Query Match 5.1%; Score 6; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 139 LAALAA 144

RESULT 274

S64789
hypothetical protein YLL038c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein L0907

C:Species: Saccharomyces cerevisiae
C:Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C:Accession: S64789
R:Duisterhoef, A.; Floeth, M.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.
submitted to the Protein Sequence Database, May 1996

A:Reference number: S64775

A:Accession: S64789

A:Molecule type: DNA

A:Residues: 1-247 <DUR>

A:Cross-references: UNIPROT:Q07872; UNIPARC:UPI000006C7F6; EMBL:Z73143; NID:gi360228; PI

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:ENT4; MIPS:YLL038C

A:Cross-references: SGD:S0003961

A:Map position: 12L

Query Match 5.1%; Score 6; DB 2; Length 247;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGLLQ 91

|||||

Db 132 VLGLLQ 137

RESULT 275

S22800

terminase gpM - phage P2

C:Species: phage P2

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

C:Accession: S22800; S16414

R:Linderoth, N.A.; Ziermann, R.; Haggard-Ljungquist, E.; Christie, G.E.; Calendar, R.

Nucleic Acids Res. 19, 7207-7214, 1991

A:Title: Nucleotide sequence of the DNA packaging and capsid synthesis genes of bacterio

A:Reference number: S22796; MUID:92115571; PMID:1837355

A:Accession: S22800

A:Molecule type: DNA

A:Residues: 1-247 <LIN>

A:Cross-references: UNIPROT:P25476; UNIPARC:UPI00000138CE2; EMBL:X61229; NID:gi515141; PID:

C:Genetics:

A:Gene: M

C:Keywords: capsid assembly; DNA binding

Query Match 5.1%; Score 6; DB 2; Length 247;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 LQRATQ 95

|||||

Db 190 LQRATQ 195

RESULT 276

T46903

hypothetical protein DKFZp761A0712.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004

C:Accession: T46903

R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24134

A:Accession: T46903

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-248 <AAA>

A:Cross-references: UNIPROT:Q9NSR5; UNIPARC:UPI00000725AF; EMBL:AL157451

A:Experimental source: adult amygdala; clone DKFZp761A0712

C:Genetics:

A>Note: DXFP761A0712.1
C:Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphatase

Query Match 5.1%; Score 6; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWIV 37
|||||
Db 71 GCWIV 76

RESULT 277
E70600
hypothetical protein RV3908 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
C:Accession: E70600
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Suleston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Accession: E70600
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-248 <COL>
A:Cross-references: UNIPROT:O05437; UNIPARC:UPI00000C1592; GB:Z94121; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3908
F:98-135/Domain: mutT domain homology <MUTT>

Query Match 5.1%; Score 6; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GHIELG 43
|||||
Db 103 GHIELG 108

RESULT 278
T33034
hypothetical protein T22B2.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 05-Oct-2004
C:Accession: T33034
R:Bentley, D.
submitted to the EMBL Data Library, February 1998
A:Description: The sequence of C. elegans cosmid T22B2.
A:Reference number: Z21267
A:Accession: T33034
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-248 <BEN>
A:Cross-references: UNIPROT:O45189; UNIPARC:UPI000007E471; EMBL:AF047662; PIDN:AAC04442.
A:Experimental source: strain Bristol N2; clone T22B2
C:Genetics:
A:Gene: CESP:T22B2.4
A:Map position: X
A:Introns: 28/1; 80/3; 124/1; 147/2; 201/1

Query Match 5.1%; Score 6; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 122 LAALAA 127

RESULT 279
F71311
probable triosephosphate isomerase (tpi) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Oct-2004
C:Accession: F71311
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwilierson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDermott, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: F71311
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-249 <COL>
A:Cross-references: UNIPROT:O83548; UNIPARC:UPI0000137238; GB:AE001229; GB:AE000520; NID:9
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0537
C:Superfamily: Triose-phosphate isomerase

Query Match 5.1%; Score 6; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VVIVGH 39
|||||
Db 89 VVIVGH 94

RESULT 280
T35724
cobalt transport integral membrane protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35724
R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, January 1998
A:Reference number: Z21548
A:Accession: T35724
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-249 <MUR>
A:Cross-references: UNIPROT:O54188; UNIPARC:UPI00000DABBF; EMBL:AL021411; PIDN:CAA16215
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: cbiQ; SCOEDB:SC7H1.29c
C:Superfamily: cobalt transport protein Q homolog

Query Match 5.1%; Score 6; DB 2; Length 249;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 234 LAALAA 239

RESULT 281
AE1288
methionine aminopeptidases homolog lmo1709 [imported] - Listeria monocytogenes (strain 1080)
C:Species: Listeria monocytogenes
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE1288
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madsen, E.; Maitournam, A.; Mouton, R.; Nuss, B.; Ok, C.; Schluter, T.; Simoes, N.; Tiersch, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A. Science 294, 849-852, 2001
A:Authors: Krest, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madsen, E.; Maitournam, A.; Mouton, R.; Nuss, B.; Ok, C.; Schluter, T.; Simoes, N.; Tiersch, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A. Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AB1288

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-252 <GLA>

A;Cross-references: UNIPROT:Q8Y6H5; UNIPARC:UPI00000554C6; GB:NC_003210; PIDN:CAC99787.1

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo1709

C;Superfamily: Escherichia coli methionyl aminopeptidase

Query Match 5.1%; Score 6; DB 2; Length 252;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 IEQAQV 77

Db 134 IEQAQV 139

RESULT 282

AD1660 methionine aminopeptidases homolog lml1821 [imported] - Listeria innocua (strain Clp112

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C;Accession: AD1660

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,

D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluecker, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AD1660

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-252 <GLA>

A;Cross-references: UNIPROT:Q92AU8; UNIPARC:UPI00000CC690; GB:AL592022; PIDN:CAC97052.1;

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lml1821

C;Superfamily: Escherichia coli methionyl aminopeptidase

Query Match 5.1%; Score 6; DB 2; Length 252;

Best Local Similarity 100.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 IEQAQV 77

Db 134 IEQAQV 139

RESULT 283

AB8955

Protein K04Fl.1 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: AB8955

R;anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog

A;Reference number: A75000; MUID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_ele

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: AB8955

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-253 <STO>

A;Cross-references: UNIPROT:Q9TXL6; UNIPARC:UPI000007C0A4; GB:chr_V; PIDN:AAC78175.1; PI

C;Genetics:

A;Gene: K04Fl.1

A;Map position: 5

Query Match 5.1%; Score 6; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24

Db 42 GVLAAAL 47

RESULT 284

B33954

hypothetical protein B - Streptomyces lividans insertion sequence IS493

C;Species: Streptomyces lividans

C;Date: 09-Mar-1990 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004

C;Accession: B33954

R;Solenberg, P.J.; Burgett, S.G.

J. Bacteriol. 171, 4807-4813, 1989

A;Title: Method for selection of transposable DNA and characterization of a new insertio

A;Reference number: A33954; MUID:89359113; PMID:2549001

A;Accession: B33954

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-254 <SOL>

A;Cross-references: UNIPROT:Q48342; UNIPARC:UPI00000A85D8; GB:M28508; NID:gl707869; PIDN:

C;Genetics:

A;Mobile element: insertion sequence IS493

A;Start codon: GTG

C;Superfamily: transposase for insertion sequence element IS112

Query Match 5.1%; Score 6; DB 2; Length 254;

Best Local Similarity 100.0%; Pred. No. 4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAAL 25

Db 79 VLAAAL 84

RESULT 285

AB1099

1-acyl-sn-glycerol-3-phosphate acyltransferase NMB1294 [imported] - Neisseria meningitid

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 16-Aug-2004

C;Accession: AB1099

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: AB1000; MUID:20175755; PMID:10710307

A;Accession: AB1099

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-255 <TEP>

A;Cross-references: UNIPROT:Q9JZ47; UNIPARC:UPI0000131C4A; GB:AE002478; GB:AE002098; NID:

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB1294

C;Superfamily: 1-acyl-sn-glycerol-3-phosphate acyltransferase

Query Match 5.1%; Score 6; DB 2; Length 255;

Best Local Similarity 100.0%; Pred. No. 4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24

Db 50 GVLAAAL 55

RESULT 286

AB1842

1-acylglycerol-3-phosphate O-acyltransferase (EC 2.3.1.51) NMA1504 [imported] - Neisseria
 C:Species: Neisseria meningitidis
 C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 16-Aug-2004
 C:Accession: A81842
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: A81842
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-255 <PAR>
 A:Cross-references: UNIPROT:Q9JU41; UNIPARC:UPI0000131C49; GB:AL162756; GB:AL157959; NID
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: plnC; NMA1504
 C:Superfamily: 1-acyl-an-glycerol-3-phosphate acyltransferase
 C:Keywords: acyltransferase; coenzyme A

Query Match 5.1%; Score 6; DB 2; Length 255;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAL 24
 |||||
 Db 50 GVLAL 55

RESULT 287
 S22363
 gufa protein homolog - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: S22363; F65091
 R:Yang, T.P.; Depew, R.E.
 Blochim. Biophys. Acta 1130, 227-228, 1992
 A:Title: Nucleotide sequence of a region duplicated in Escherichia coli toc mutants.
 A:Reference number: S22360; MUID:92223101; PMID:1314093
 A:Accession: S22363
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-257 <YAN>
 A:Cross-references: UNIPROT:P24198; UNIPARC:UPI0000046C4F; EMBL:M77129; NID:g146676; PID
 A:Experimental source: strain RED44
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: F65091
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-257 <BLAT>
 A:Cross-references: UNIPARC:UPI0000046C4F; GB:AE000386; GB:U00096; NID:g2367187; PIDN:AP
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ygiE
 C:Superfamily: gufa protein

Query Match 5.1%; Score 6; DB 1; Length 257;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 183 LGGVLA 188

RESULT 288
 AF0890
 probable membrane protein STV3368 [imported] - Salmonella enterica subsp. enterica serov
 C:Species: Salmonella enterica subsp. enterica serovar Typhi

A:Note: this species has also been called Salmonella typhi
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AF0890
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AF0890
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <PAR>
 A:Cross-references: UNIPARC:UPI000005A49A; GB:AL513382; PIDN:CAD07716.1; PID:g16504268;
 C:Genetics:
 A:Gene: STV3368
 C:Superfamily: gufa protein

Query Match 5.1%; Score 6; DB 2; Length 257;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 183 LGGVLA 188

RESULT 289
 H9119
 gufa protein homolog - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
 C:Species: Escherichia coli
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: H9119
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H9119
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <HAY>
 A:Cross-references: UNIPROT:P24198; UNIPARC:UPI0000046C4F; GB:BA000007; PIDN:BA037351.1;
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs3928
 C:Superfamily: gufa protein

Query Match 5.1%; Score 6; DB 2; Length 257;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 183 LGGVLA 188

RESULT 290
 G85964
 gufa protein homolog - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C:Accession: G85964
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: G85964
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <STO>

A;Cross-references: UNIPROT:P24198; UNIPARC:UPI0000046C4P; GB:AE005174; NID:gi2517621; E
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ygiS
C;Superfamily: gufA protein

Query Match 5.1%; Score 6; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 183 LGGVLA 188

RESULT 291
S16865
Gene f41 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Oct-2004
C;Accession: S16865
R;Kato, K.
submitted to the EMBL Data Library, August 1991
A;Description: Sequence analysis of twenty mouse brain cDNA clones selected by specific
A;Reference number: S16416
A;Accession: S16865
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-257 <KAT>
A;Cross-references: UNIPROT:P28659; UNIPARC:UPI000016CD44; EMBL:X61451; NID:g50941; PIDN
P;1-49/Domain: ribonucleoprotein repeat homology #status atypical <RRM1>

Query Match 5.1%; Score 6; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 138 LAALAA 143

RESULT 292
C87287
Inositol monophosphatase family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87287
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87287
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-259 <STO>
A;Cross-references: UNIPROT:Q9ABC2; UNIPARC:UPI000006PFD4; GB:AE005673; NID:gi3421451; E
C;Genetics:
A;Gene: CC0308
C;Superfamily: Aquifex aeolicus cysQ protein

Query Match 5.1%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 37 VGHTEL 42
|||||
Db 253 VGHTEL 258

RESULT 293
C70615
probable chromosome partitioning ATPase, ParA family - Deinococcus radiodurans (strain R
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: E75592
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

hypothetical protein Rv0128 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: C70615
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: C70615
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-259 <COL>
A;Cross-references: UNIPROT:P96805; UNIPARC:UPI00000D0F5E; GB:Z92770; GB:AL123456; NID:g
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv0128
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0128

Query Match 5.1%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
|||||
Db 15 VLAALA 20

RESULT 294
S77575
oligopeptide transport protein homolog ophD - Agrobacterium tumefaciens plasmid pTiR10
N;Alternate names: oligopeptide permease homolog ophD
C;Species: Agrobacterium tumefaciens
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Oct-2004
C;Accession: S77575
R;Fuqua, C.; Winans, S.C.
Mol. Microbiol. 20, 1199-1210, 1996
A;Title: Localization of Occr-activated and TraR-activated promoters that express two AB
A;Reference number: S77571; MUID:96405643; PMID:8809772
A;Accession: S77575
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-259 <FOU>
A;Cross-references: UNIPROT:Q44375; UNIPARC:UPI00000B7956; EMBL:U48718; NID:gi215729; P
A;Experimental source: strain RS10
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1996
C;Genetics:
A;Gene: ophD
A;Genome: plasmid pTiR10
A;Keywords: ATP; nucleotide binding; oligopeptide transport; P-loop
P;26-231/Domain: ATP-binding cassette homology <ABC>
P;43-50/Region: nucleotide-binding motif A (P-loop)

Query Match 5.1%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
|||||
Db 36 GKVLGL 41

RESULT 295
E75592
probable chromosome partitioning ATPase, ParA family - Deinococcus radiodurans (strain R
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: E75592
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: E75592
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-260 <WHI>
 A:Cross-references: UNIPROT:Q9RZB8; UNIPARC:UPI0000003B7D; GB:AE001862; GB:AE001825; NID
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DRA0001
 A:Map position: 2
 C:Superfamily: regulatory protein spo0J

Query Match 5.1%; Score 6; DB 2; Length 260;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 136 LAALAA 141

RESULT 296

D70871
 hypothetical protein Rv1457c - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: D70871
 R: Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtrooyd, S.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Scares, R.; Suleston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: D70871
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-261 <COL>
 A:Cross-references: UNIPROT:O86349; UNIPARC:UPI000000605A; GB:AL021184; GB:AL123456; NID
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: Rv1457c
 C:Superfamily: Streptomyces peucetius daunorubicin resistance protein

Query Match 5.1%; Score 6; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 249 LAALAA 254

RESULT 297

T43579
 type III secretion protein yscT - *Yersinia pestis* plasmid pCD1
 N:Alternate names: translocation protein T homolog
 C:Species: *Yersinia pestis*
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 05-Oct-2004
 C:Accession: T43579; T42874
 R:Hu, P.; Elliott, J.; McCready, P.; Skowronski, E.; Garnes, J.; Kobayashi, A.; Brubaker
 J. Bacteriol. 180, 5192-5202, 1998
 A:Title: Structural organization of virulence-associated plasmids of *Yersinia pestis*.
 A:Reference number: 225578; MUID:98422474; PMID:9748454
 A:Accession: T43579
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-261 <HUP>
 A:Cross-references: UNIPROT:P40299; UNIPARC:UPI000013COA1; EMBL:AF053946; NID:92996222;
 A:Experimental source: strain KIM
 R: Perry, R. D.; Straley, S. C.; Fetherston, J. D.; Rose, D. J.; Gregor, J.; Blattner, F. R.

Infect. Immun. 66, 4611-4623, 1998
 A:Title: DNA sequencing and analysis of the low-Ca2+-response plasmid pCD1 of *Yersinia*
 A:Reference number: 222273; MUID:98427122; PMID:9746557
 A:Accession: T42874
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-261 <PER>
 A:Cross-references: UNIPARC:UPI000013COA1; EMBL:AF074612; NID:G3822037; PIDN:AAC69785.1
 A:Experimental source: strain KIM5
 C:Genetics:
 A:Gene: yscT
 A:Genome: plasmid pCD1
 C:Superfamily: Yop proteins translocation protein T

Query Match 5.1%; Score 6; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||
 Db 34 LLGGVL 39

RESULT 298

B85651
 hypothetical protein Z1555 [imported] - *Escherichia coli* (strain O157:H7, substrain EDL
 C:Species: *Escherichia coli*
 C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 15-Mar-2004
 C:Accession: B85651
 R: Perna, N. T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J. D.; Rose, D. J.; Mayhew
 Miller, L.; Grobeck, E. J.; Davis, N. W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B85651
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-262 <STO>
 A:Cross-references: UNIPARC:UPI0000165761; GB:AE005174; NID:G12514426; PIDN:AAG55670.1;
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: Z1555
 C:Superfamily: uncharacterized conserved protein

Query Match 5.1%; Score 6; DB 2; Length 262;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 KEVLQ 57
 |||||
 Db 204 KEVLQ 209

RESULT 299

G90790
 hypothetical protein ECs1295 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
 C:Species: *Escherichia coli*
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 15-Mar-2004
 C:Accession: G90790
 R: Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C. G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: G90790
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-262 <HAY>
 A:Cross-references: UNIPARC:UPI000000D06C8; GB:BA000007; PIDN:BNB34718.1; PID:G13360755;
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: ECs1295
 C:Superfamily: uncharacterized conserved protein

Query Match 5.1k; Score 6; DB 2; Length 262;
Best Local Similarity 100.0k; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 52 KEVLYQ 57
|||||
Db 204 KEVLYQ 209

RESULT 300
H75377
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: H75377
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75377
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <WHI>
A:Cross-references: UNIPROT:Q9RU13; UNIPARC:UPI00000C196A; GB:AE002002; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1583
A:Map position: 1

Query Match 5.1k; Score 6; DB 2; Length 262;
Best Local Similarity 100.0k; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
|||||
Db 144 GGVLA 149

RESULT 301
AD2819
conserved hypothetical protein Atu1975 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AD2819
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD2819
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-262 <KUR>
A:Cross-references: UNIPROT:Q8UD21; UNIPARC:UPI000016467D; GB:AE008688; PIDN:AAL42970.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu1975
A:Map position: circular chromosome

Query Match 5.1k; Score 6; DB 2; Length 262;
Best Local Similarity 100.0k; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 42 LGGVLA 47

RESULT 302
H83170
conserved hypothetical protein PA3797 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2004
C:Accession: H83170
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83170
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <STO>
A:Cross-references: UNIPROT:Q9HXK0; UNIPARC:UPI00000C5B19; GB:AE004798; GB:AE004091; NID
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA3797
C:Superfamily: nitrilase (carbon-nitrogen hydrolase)

Query Match 5.1k; Score 6; DB 2; Length 264;
Best Local Similarity 100.0k; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
|||||
Db 242 AALAA 247

RESULT 303
S72778
hypothetical protein B1496 F2.59 - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S72778; T11016
R:Smith, D.R.; Robinson, K.
submitted to the EMBL Data Library, November 1993
A:Description: Mycobacterium leprae cosmid B1496.
A:Reference number: S72695
A:Accession: S72778
A:Molecule type: DNA
A:Residues: 1-265 <SMI>
A:Cross-references: UNIPROT:Q49703; UNIPARC:UPI00000D437C; EMBL:U00013; NID:g466868; PID
R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z16918
A:Accession: T11016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-265 <PAR>
A:Cross-references: UNIPARC:UPI00000D437C; EMBL:Z99125; NID:g2398683; PIDN:CAB16176.1; P
C:Genetics:
A:Gene: MLCU536.32
A:Start codon: GTG
C:Superfamily: Streptomyces peucetius daunorubicin resistance protein

Query Match 5.1k; Score 6; DB 2; Length 265;
Best Local Similarity 100.0k; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 253 LAALAA 258

RESULT 304
T31217
transcription regulator homolog - Sphingomonas aromaticivorans plasmid pML1
C:Species: Sphingomonas aromaticivorans
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004

C:Accession: T31217
R:Romane, M.P.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.;
submitted to the EMBL Data Library, July 1998
A:Description: Complete sequence of a 184 kb catabolic plasmid from *Sphingomonas aromati*
A:Reference number: Z20992
A:Accession: T31217
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-266 <ROM>
A:Cross-references: UNIPROT:O85926; UNIPARC:UPI000005C9B2; EMBL:AF079317; NID:G3378261;
C:Genetics:
A:Genome: plasmid pNL1
A:Note: orf758

Query Match 5.1%; Score 6; DB 2; Length 266;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 107 LAALAA 112

RESULT 305
G97450
ABC transporter, ATP-binding protein [imported] - *Agrobacterium tumefaciens* (strain C58,
C:Species: *Agrobacterium tumefaciens*
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97450
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-267 <KUR>
A:Cross-references: UNIPROT:Q8UHD1; UNIPARC:UPI00000D18F9; GB:AE007869; PIDN:AAK86560.1;
C:Genetics:
A:Gene: AGR_C_1362
A:Map position: circular chromosome
C:Superfamily: syphilis spirochete probable ABC transporter nara; ATP-binding cassette h

Query Match 5.1%; Score 6; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 182 LAALAA 187

RESULT 306
S74415
hypothetical protein al10688 - *Synechocystis* sp. (strain PCC 6803)
C:Species: *Synechocystis* sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S74415
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74415
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-267 <KAN>
A:Cross-references: UNIPROT:Q55191; UNIPARC:UPI00000C0BDF; EMBL:D64001; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 5.1%; Score 6; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 256 LAALAA 261

RESULT 308
S62760
cytochrome-c oxidase (EC 1.9.3.1) chain III - red alga (*Cyanidium caldarium*) mitochondrion
C:Species: mitochondrion *Cyanidium caldarium*
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S62760
R:Viehmann, S.
submitted to the EMBL Data Library, March 1995
A:Reference number: S62757
A:Accession: S62760
A:Molecule type: DNA
A:Residues: 1-270 <VIE>
A:Cross-references: UNIPROT:P48873; UNIPARC:UPI00001280BA; EMBL:Z48930; NID:G791089; PFI
C:Genetics:
A:Gene: coliII
A:Genome: mitochondrion
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
F:7-21/Domain: mitochondrial matrix #status predicted <MM1>
F:22-40/Domain: transmembrane #status predicted <TM01>
F:41-48/Domain: intracristal #status predicted <ITC1>
F:49-74/Domain: transmembrane #status predicted <TM02>
F:75-80/Domain: mitochondrial matrix #status predicted <MM2>
F:81-113/Domain: transmembrane #status predicted <TM03>
F:114-136/Domain: intracristal #status predicted <ITC2>
F:137-160/Domain: transmembrane #status predicted <TM04>
F:161-163/Domain: mitochondrial matrix #status predicted <MM3>

F;164-191/Domain: transmembrane #status predicted <TM05>
F;192-198/Domain: intracistal #status predicted <ITC3>
F;199-231/Domain: transmembrane #status predicted <TM06>
F;232-240/Domain: mitochondrial matrix #status predicted <MM4>
F;241-264/Domain: transmembrane #status predicted <TM07>
F;265-269/Domain: intracistal #status predicted <ITC4>

Query Match 5.1%; Score 6; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVIL 21
| | | | |
Db 34 LGGVIL 39

RESULT 309

H70690
hypothetical protein Rv2813 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70690
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70690
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-270 <COL>
A;Cross-references: UNIPROT:P71638; UNIPARC:UPI00000D5BB2; GB:Z81331; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv2813

Query Match 5.1%; Score 6; DB 2; Length 270;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 176 GVLAAL 181

RESULT 310

Tl1590
hypothetical protein SPAC17C9.10 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: Tl1590; T52026
R;Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
A;Reference number: Z17295
A;Accession: Tl1590
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-271 <BAR>
A;Cross-references: UNIPROT:Q10482; UNIPARC:UPI0000136110; EMBL:Z73099; NID:g1314152
A;Experimental source: strain 972h (-)
R;Kyung Sook, K.S.
A;Reference number: Z25907
A;Accession: T52026
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-271 <KYU>
A;Cross-references: UNIPARC:UPI0000136110; EMBL:L49134; PIDN:AAD45180.1
A;Experimental source: strain ED665h (-)
C;Genetics:
A;Map position: IR

A;Note: SPAC17C9.10: stml
C;Superfamily: Saccharomyces probable membrane protein YBR147w
C;Keywords: transmembrane protein

Query Match 5.1%; Score 6; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIV 37
| | | | |
Db 148 GCVVIV 153

RESULT 311

T37222
probable secreted protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T37222
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1998
A;Reference number: Z21615
A;Accession: T37222
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-271 <OLI>
A;Cross-references: UNIPROT:O86589; UNIPARC:UPI00000DADB7; EMBL:AL031514; PIDN:CAA20600.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOBDB:SC2H4.06C

Query Match 5.1%; Score 6; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 245 VLAALA 250

RESULT 312

F70703
probable uspE protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: F70703
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: F70703
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-274 <COL>
A;Cross-references: UNIPROT:P71895; UNIPARC:UPI00000D11F0; GB:Z79702; GB:AL123456; NID:9
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: uspE
C;Superfamily: maltose transport protein malG

Query Match 5.1%; Score 6; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 193 VLAALA 198

RESULT 313

T35062
Probable citrate lyase beta subunit - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 24-Nov-2003
C:Accession: T35062
R:Seeger, K.J.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z1567
A:Accession: T35062
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-274 <SEE>
A:CROSS-references: UNIPARC:UPI00000DB2A2; EMBL:AL096884; PIDN:CAB51425.1; GSPDB:GN00070
C:Experimental source: strain A3(2)
C:Genetics:
C:Superfamily: citrate lyase, subunit beta

Query Match 5.1%; Score 6; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 78 LAALAA 83

RESULT 314
E82179
ABC transporter, periplasmic substrate-binding protein-related protein VC1598 [imported]
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82179
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E.
l.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82179
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-274 <HEI>
A:CROSS-references: UNIPROT:Q9KXN8; UNIPARC:UPI00000C303D; GB:AE004237; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1598
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 164 LGGVLA 169

RESULT 315
D87407
conserved hypothetical protein CC1275 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: D87407
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87407
A:Status: preliminary
A:Molecule type: DNA

Query Match 5.1%; Score 6; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAAAL 25
|||||
DB 161 VLAAAL 166

RESULT 318

A:Residues: 1-276 <STO>
A:CROSS-references: UNIPROT:Q9A8S6; UNIPARC:UPI00000C731F; GB:AE005673; NID:gl3422608; F
C:Genetics:
A:Gene: CC1275

Query Match 5.1%; Score 6; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 170 LAALAA 175

RESULT 316

B83721
oligopeptide ABC transporter (ATP-binding protein) BH0570 [imported] - Bacillus halodurans
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: B83721
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83721
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <STO>
A:CROSS-references: UNIPROT:Q9KFB5; UNIPARC:UPI00000C3901; GB:AP001509; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0570

Query Match 5.1%; Score 6; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
|||||
DB 38 GKVLGL 43

RESULT 317

H84314
cytochrome aa3 controlling protein [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: H84314
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: H84314
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-277 <STO>
A:CROSS-references: UNIPROT:Q9HP13; UNIPARC:UPI0000063931; GB:AE004437; NID:gl0581096; F
C:Genetics:
A:Gene: ccp

Query Match 5.1%; Score 6; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAAAL 25
|||||
DB 161 VLAAAL 166

RESULT 318

H95113
conserved hypothetical protein SP0986 [imported] - Streptococcus pneumoniae (strain TIGR)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: H95113
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapfel,
neon, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Lofthus, B.J.; Yang, P.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: H95113
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <KUR>
A;Cross-references: UNIPROT:Q97R48; UNIPARC:UPI0000051683; GB:AE005672; PIDN:AAK75105.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0986

Query Match 5.1%; Score 6; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 25 GGVLA 30

RESULT 319
A97983
conserved hypothetical protein spr0889 [imported] - Streptococcus pneumoniae (strain R6)
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: A97983
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Ballido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: A97983
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <KUR>
A;Cross-references: UNIPROT:Q8DQ19; UNIPARC:UPI000008354B; GB:AE007317; PIDN:AAK99693.1;
C;Genetics:
A;Gene: spr0889

Query Match 5.1%; Score 6; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 25 GGVLA 30

RESULT 320
C95932
probable sugar uptake ABC transporter permease protein SMB21219 [imported] - Sinorhizob
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95932
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: C95932
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-281 <KUR>
A;Cross-references: UNIPROT:Q92V10; UNIPARC:UPI000000CB62C; GB:AL591985; PIDN:CAC49123.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, F.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB21219
A;Genome: plasmid
C;Superfamily: maltose transport protein malG

Query Match 5.1%; Score 6; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAA 24
Db 252 GVLAA 257

RESULT 321
T05522
hypothetical protein F13M23.150 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T05522
R;Bevan, M.; Wedler, H.; Wedler, E.; Wambutt, R.; Hoheisel, J.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15419
A;Accession: T05522
A;Molecule type: DNA
A;Residues: 1-281 <BEV>
A;Cross-references: UNIPROT:Q98W25; UNIPARC:UPI000000A989E; EMBL:AL035523
A;Experimental source: cultivar Columbia; BAC clone F13M23
C;Genetics:
A;Map position: 4
A;Introns: 16/1; 28/2; 99/3; 153/3; 193/3
A;Note: F13M23.150

Query Match 5.1%; Score 6; DB 2; Length 281;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGL 90
Db 104 KVLGL 109

RESULT 322
C64241
hypothetical protein MG373 - Mycoplasma genitalium
C;Species: Mycoplasma genitalium
C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 05-Oct-2004
C;Accession: C64241
R;Praeger, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.;
M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.
Science 270, 397-403, 1995
A;Title: The minimal gene complement of Mycoplasma genitalium.
A;Reference number: A64200; MUID:96026346; PMID:7569993
A;Accession: C64241
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-281 <TIGR>
A;Cross-references: UNIPROT:P47613; UNIPARC:UPI00001396DE; GB:U39722; GB:L43967; NID:910
A;Experimental source: strain G-37
C;Genetics:

A:Genetic code: SGC3
 C:Superfamily: uncharacterized conserved protein

Query Match 5.1%; Score 6; DB 2; Length 281;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
 DB 41 GKVLGL 46

RESULT 323
 H75250
 OxA-related protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: H75250
 R:White, O.; Eisen, J.A.; Heidelberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: H75250
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-282 <WHI>
 A:Cross-references: UNIPROT:Q9RR64; UNIPARC:UPI000003DFPC; GB:AE002092; GB:AE000513; NID
 A:Experimental source: strain R1
 A:Genetics:
 A:Gene: DR2633
 A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 282;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
 DB 151 VLAALA 156

RESULT 324
 A64999
 hypothetical protein b2275 - Escherichia coli (strain K-12)
 C:Species: Escherichia coli
 C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C:Accession: A64999
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A:Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:197426617; PMID:9278503
 A:Accession: A64999
 A>Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-283 <BLAT>
 A:Cross-references: UNIPROT:P76486; UNIPARC:UPI000013ADE3; GB:AE000317; GB:U00096; NID:9
 A:Experimental source: strain K-12, substrain MG1655
 C:Superfamily: Escherichia coli hypothetical protein b2275

Query Match 5.1%; Score 6; DB 2; Length 283;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LIQRAT 94
 DB 123 LIQRAT 128

RESULT 325
 H69369

branched-chain amino acid ABC transporter, permease protein (braE-3) homolog - Archaeog
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: H69369
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.P.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
 Nature 390, 364-370, 1997
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
 A:Reference number: A69250; MUID:198049343; PMID:9389475
 A:Accession: H69369
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-285 <KLE>
 A:Cross-references: UNIPROT:Q29302; UNIPARC:UPI0000056BEP; GB:AE001038; GB:AE000782; N
 C:Superfamily: leucine transport protein livH

Query Match 5.1%; Score 6; DB 2; Length 285;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29
 DB 63 LAAYCL 68

RESULT 326
 A87534
 carboxylesterase family protein [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: A87534
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolo
 n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: A87534
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-289 <STO>
 A:Cross-references: UNIPROT:Q9A5Z9; UNIPARC:UPI00000C769B; GB:AE005673; NID:gl3423817;
 C:Genetics:
 A:Gene: CC2298

Query Match 5.1%; Score 6; DB 2; Length 289;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 DB 12 LAALAA 17

RESULT 327
 AB32235
 nitrilotriacetate monooxygenase Atu6060 [imported] - Agrobacterium tumefaciens (strain C
 C:Species: Agrobacterium tumefaciens
 C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AB32235
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 i; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AB3235
 A>Status: preliminary

```
A:Molecule type: DNA
A:Residues: 1-291 <KUR>
A:Cross-references: UNIPROT:Q8U675; UNIPARC:UPI0000167CF5; GB:AE008690; PIDN:AAL46296.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu6060
A:Genome: plasmid

Query Match          5.1%; Score 6; DB 2; Length 291;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
    |||||
Db 52 LAALAA 57

RESULT 328
T47163
hypothetical protein DKFZp762E1312.1 - human
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47163
R:Ansorge, W.; Winkler, U.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24375
A:Accession: T47163
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-293 <AAA>
A:Cross-references: UNIPROT:Q9NSL8; UNIPARC:UPI0000071748; EMBL:AL162048
A:Experimental source: adult melanoma (MeWo cell line); clone DKFZp762E1312
C:Genetics:
A:Note: DKFZp762E1312.1

Query Match          5.1%; Score 6; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEV 54
    |||||
Db 109 VPDKEV 114

RESULT 329
A87509
hypothetical protein CC2098 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87509
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87509
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <STO>
A:Cross-references: UNIPROT:Q9AGJ5; UNIPARC:UPI000012F0B5; GB:AE005673; NID:g13423583; E
C:Genetics:
A:Gene: CC2098
C:Superfamily: delta(2)-isopentenylpyrophosphate transferase

Query Match          5.1%; Score 6; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIV 49
    |||||
Db 97 GKPAIV 102
```

RESULT 330

```
E97597
hypothetical protein AGR_C_3597 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: E97597
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: E97597
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KUR>
A:Cross-references: UNIPROT:Q8UDZ1; UNIPARC:UPI000000D1D32; GB:AE007869; PIDN:AAK87734.1;
C:Genetics:
A:Gene: AGR_C_3597
A:Map position: circular chromosome
```

```
Query Match          5.1%; Score 6; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 17 LGGVLA 22
    |||||
Db 75 LGGVLA 80
```

RESULT 331

```
T05110
hypothetical protein F28M20.210 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05110
R:Bevan, M.; Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Hoheisel, J.; Mewes,
submitted to the Protein Sequence Database, November 1998
A:Reference number: Z15398
A:Accession: T05110
A:Molecule type: DNA
A:Residues: 1-296 <BEV>
A:Cross-references: UNIPROT:Q9SB76; UNIPARC:UPI000000A48D9; EMBL:AL031004
A:Experimental source: cultivar Columbia; BAC clone F28M20
C:Genetics:
A:Map position: 4
A:Introns: 22/3; 104/3; 141/3; 194/3; 242/3
A:Note: F28M20.210
```

```
Query Match          5.1%; Score 6; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 15 VLLGGV 20
    |||||
Db 247 VLLGGV 252
```

RESULT 332

```
GB3056
conserved hypothetical protein PA4717 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: GB3056
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Latbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: GB3056
A:Status: preliminary
A:Molecule type: DNA
```

```
A:Residues: 1-297 <STO>
A:Cross-references: UNIPROT:Q9HV81; UNIPARC:UPI000000C5DDF; GB:AE004885; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4717

Query Match          5.1%; Score 6; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 ATQOQA 98
Db 70 ATQOQA 75

RESULT 333
AG2955
hypothetical protein Atu3245 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AG2955
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClellan,
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG2955
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-297 <KUR>
A:Cross-references: UNIPROT:Q8UAX2; UNIPARC:UPI00000D2128; GB:AE008689; PIDN:AAL44061.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3245
A:Map position: linear chromosome

Query Match          5.1%; Score 6; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 204 LAALAA 209

RESULT 334
F98327
hypothetical protein AGR_L_3137 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: F98327
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: F98327
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-297 <KUR>
A:Cross-references: UNIPROT:Q8UAX2; UNIPARC:UPI00000D2128; GB:AE007870; PIDN:AAK90144.1;
C:Genetics:
A:Gene: AGR_L_3137
A:Map position: linear chromosome

Query Match          5.1%; Score 6; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
```

```
Db 204 LAALAA 209

RESULT 335
E84310
hypothetical protein Vng1572c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84310
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: E84310
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <STO>
A:Cross-references: UNIPROT:Q9HPL8; UNIPARC:UPI0000063910; GB:AE004437; NID:gl0581055;
C:Genetics:
A:Gene: VNG1572C
C:Superfamily: uncharacterized conserved protein MJ1598

Query Match          5.1%; Score 6; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 117 GVLAAL 122

RESULT 336
B83100
inhibitor of chromosome initiation IciA PA4363 [imported] - Pseudomonas aeruginosa (str
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83100
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: B83100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-300 <STO>
A:Cross-references: UNIPROT:Q9HW38; UNIPARC:UPI000012D167; GB:AE004852; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: iciA; PA4363

Query Match          5.1%; Score 6; DB 2; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 9 LAALAA 14

RESULT 337
S57531
transcription regulator - Acinetobacter calcoaceticus
C:Species: Acinetobacter calcoaceticus
C>Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S57531
R:Kok, R.G.; Bart, A.; Hellingwerf, K.J.
submitted to the EMBL Data Library, June 1995
```

A;Description: Characterization of the estBR operon of Acinetobacter calcoaceticus BD413
A;Reference number: S57529
A;Accession: S57531
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-301 <KOK>
A;Cross-references: UNIPROT:P52667; UNIPARC:UPI000012A20B; EMBL:X88895; NID:g887382; PID
C;Superfamily: regulatory protein ilvY

Query Match 5.1%; Score 6; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVIL 21
|||||
Db 46 LGGVIL 51

RESULT 338
C72616
hypothetical protein APE1388 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: C72616
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: C72616
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-301 <KAW>
A;Cross-references: UNIPROT:Q9YC63; UNIPARC:UPI000005DF1E; DDBJ:AP000061; NID:g5104821;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1388
C;Superfamily: Aeropyrum pernix hypothetical protein APE1388

Query Match 5.1%; Score 6; DB 2; Length 301;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 150 LAALAA 155

RESULT 339
T35246
probable integral membrane protein - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T35246
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21573
A;Accession: T35246
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-302 <OLI>
A;Cross-references: UNIPROT:Q9XYM7; UNIPARC:UPI00000DAF9C; EMBL:AL049587; PIDN:CAB40671.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC08DB:SCSF2A.04

Query Match 5.1%; Score 6; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 239 LAALAA 244

RESULT 340
S27846
hypothetical protein - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: S27846
R;Woodward, R.; Carden, M.J.; Gull, K.
submitted to the EMBL Data Library, March 1992
A;Reference number: S27846
A;Accession: S27846
A;Molecule type: mRNA
A;Residues: 1-302 <WOO>
A;Cross-references: UNIPROT:Q26784; UNIPARC:UPI0000077E3F; EMBL:M87318; NID:g162176; PID

Query Match 5.1%; Score 6; DB 2; Length 302;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQR 92
|||||
Db 6 LGLLQR 11

RESULT 341
S75782
methanol dehydrogenase regulatory protein - Synecocystis sp. (strain PCC 6803)
N;Alternate names: protein slr0835
C;Species: Synecocystis sp.
A;Variety: PCC 6803
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: S75782
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S75782
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-303 <KAN>
A;Cross-references: UNIPROT:Q55419; UNIPARC:UPI00000C10F6; EMBL:D64003; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Genetics:
A;Gene: moxR
C;Superfamily: methanol dehydrogenase regulatory protein

Query Match 5.1%; Score 6; DB 1; Length 303;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 LQRATQ 95
|||||
Db 249 LQRATQ 254

RESULT 342
G70756
hypothetical protein RV1985c - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: G70756
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: G70756
A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-303 <COL>
 A:Cross-references: UNIPROT:Q10872; UNIPARC:UPI000013B4E4; GB:Z74025; GB:AL123456; NID:9
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: RV1985c
 C:Superfamily: conserved hypothetical protein H11364

Query Match 5.1%; Score 6; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 11 LAALAA 16

RESULT 343

T36509
 probable molybdopterin-guanine dinucleotide biosynthesis protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T36509
 R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z21608
 A:Accession: T36509
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-303 <SAU>
 A:Cross-references: UNIPROT:Q9XA50; UNIPARC:UPI00000DB1FD; EMBL:AL096822; PIDN:CAB46951.
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCGD3.29c

Query Match 5.1%; Score 6; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 65 LAALAA 70

RESULT 344

DB7352
 conserved hypothetical protein CC0831 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C:Accession: DB7352
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: DB7352
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-303 <STO>
 A:Cross-references: UNIPROT:Q9A9Y0; UNIPARC:UPI00000C71AB; GB:AE005673; NID:g13422082; F
 C:Genetics:
 A:Gene: CC0831

Query Match 5.1%; Score 6; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRA 93
 |||||
 DB 185 GLLQRA 190

RESULT 345

CB2524
 pyrrhline-5-carboxylate reductase XF2712 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: CB2524
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: AB2515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: CB2524
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-305 <SIM>
 A:Cross-references: UNIPROT:Q9PA08; UNIPARC:UPI00000C2B16; GB:AE004077; GB:AE003849; NID:
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, I.
 as-Neto, B.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Proh
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.,
 , P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas
 A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF2712
 C:Superfamily: pyrrhline-5-carboxylate reductase

Query Match 5.1%; Score 6; DB 2; Length 305;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TSTWVL 16
 |||||
 DB 92 TSTWVL 97

RESULT 346

A40573
 clathrin heavy chain - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
 C:Accession: A40573
 R:Dodge, G.R.; Kovalsky, I.; McBride, O.W.; Yi, H.F.; Chu, M.; Saitta, B.; Stokes, D.G.
 Genomics 11, 174-178, 1991
 A:Title: Human clathrin heavy chain (CLTC): partial molecular cloning, expression, and
 A:Reference number: A40573; MUID:92112210; PMID:1765375
 A:Accession: A40573
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-305 <DOD>
 A:Cross-references: UNIPROT:Q00610; UNIPARC:UPI000016A6DE; GB:X55878; GB:S75467; NID:925
 C:Superfamily: clathrin heavy chain

Query Match 5.1%; Score 6; DB 2; Length 305;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRA 93
 |||||
 DB 62 GLLQRA 67

RESULT 347

F72574
 probable cytochrome C oxidase assembly factor APB1878 - Aeropyrum pernix (strain K1)

C;Species: Aetopyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: F72574
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kana Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aetopyrum pernix
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: F72574
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-305 <KAW>
A;Cross-references: UNIPROT:Q9VAR5; UNIPARC:UPI000005E10F; DDBJ:AP000062; NID:gs105244;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1878
C;Superfamily: heme O synthase

Query Match 5.1%; Score 6; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
|||||
Db 253 VLAALA 258

RESULT 348
AC2419
ornithine carbamoyltransferase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2004
C;Accession: AC2419
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC2419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-306 <KUB>
A;Cross-references: UNIPROT:Q8YMM6; UNIPARC:UPI00000CEBEC; GB:BA000019; PIDN:BA076606.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr4907
C;Superfamily: ornithine/aspartate carbamoyltransferase; aspartate/ornithine carbamoyltransferase

Query Match 5.1%; Score 6; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLL 90
|||||
Db 41 KVLGLL 46

RESULT 349
T46757
lipoprotein lmb [validated] - Streptococcus agalactiae
C;Species: Streptococcus agalactiae
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T46757
R;Spellerberg, B.; Rozdzinski, B.; Martin, S.; Weber-Heynemann, J.; Schnitzler, N.; Luetjens, Immun. 67, 871-878, 1999
A;Title: Lmb, a protein with similarities to the Lrai adhesin family, mediates attachment of Streptococcus agalactiae to human laminin
A;Reference number: Z24091; MUID:99115568; PMID:9916102
A;Accession: T46757
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-306 <SPS>
A;Cross-references: UNIPROT:Q9ZHG8; UNIPROT:Q8DZ80; UNIPARC:UPI00000B30F0; EMBL:AF062533

A;Experimental source: strain R268
C;Genetics:
A;Gene: lmb
C;Function:
A;Description: mediates the attachment of S. agalactiae to human laminin [validated], MUI
C;Superfamily: adhesin B

Query Match 5.1%; Score 6; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 53 EVLYQQ 58
|||||
Db 299 EVLYQQ 304

RESULT 350
C70112
conserved hypothetical protein BB0099 - Lyme disease spirochete
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: C70112
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White, S.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.
A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: C70112
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-307 <KLE>
A;Cross-references: UNIPROT:O51126; UNIPARC:UPI00000322B4; GB:AE001122; GB:AE000783; NID:
A;Experimental source: strain B31
C;Superfamily: conserved hypothetical protein H11714

Query Match 5.1%; Score 6; DB 1; Length 307;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 FKGVKL 87
|||||
Db 31 FKGVKL 36

RESULT 351
S22931
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - gray-crowned babbler mitoc
C;Species: mitochondrion Pomatostomus temporalis (gray-crowned babbler)
C;Date: 29-Jan-1993 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: S22931; H33285; S21611; S70405; S70404; S21612; S21613; S21614; S21615; S21
R;Edwards, S.V.; Arcander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A;Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A;Reference number: S22919; MUID:91288587; PMID:1676522
A;Accession: S22931
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <EDW>
A;Cross-references: UNIPROT:P29631; UNIPARC:UPI0000128AAB; EMBL:X60936; NID:gl3395; PIDN:
R;Kocher, T.D.; Thomas, W.K.; Meyer, A.; Edwards, S.V.; Paeabo, S.; Villablanca, F.X.; Proc. Natl. Acad. Sci. U.S.A. 86, 6196-6200, 1989
A;Title: Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing
A;Reference number: A33285; MUID:89345630; PMID:2762322
A;Accession: H33285
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 15,'A','17-94 <KOC>
A;Cross-references: UNIPARC:UPI000016D60C; GB:M25688; NID:g343668; PIDN:AAA32139.1; PID:
R;Edwards, S.V.; Wilson, A.C.
submitted to the EMBL Data Library, October 1990
A;Description: Phylogenetically informative length polymorphism and sequence variability

A;Reference number: S21611
A;Accession: S21611
A:Molecule type: DNA
A;Residues: 5-98 <ED2>
C;Cross-references: UNIPARC:UPI000016D60B; EMBL:X54900; EMBL:X54901; EMBL:X54902; EMBL:X54903; EMBL:X54891
EMBL:X54893; EMBL:X54892; EMBL:X54891
R;Edwards, S.V.; Wilson, A.C.
Genetics 126, 695-711, 1990
A>Title: Phylogenetically informative length polymorphism and sequence variability in mid
A;Reference number: S70404; MUID:91065505; PMID:1979038
A;Accession: S70405
A:Molecule type: DNA
A;Residues: 5-98 <EDF>
C;Cross-references: UNIPARC:UPI000016D60B; EMBL:X54891; EMBL:X54906; EMBL:X54892; EMBL:X54905
EMBL:X54902; EMBL:X54903; EMBL:X54904; EMBL:X54905
A;Accession: S70404
A;Status: translation not shown
A:Molecule type: DNA
A;Residues: 5-13,'I',15-98 <ED3>
C;Cross-references: UNIPARC:UPI00000922A1; EMBL:X54885; NID:g13329; PIDN:CAA38658.1; PID:
A;Experimental source: individual 23C
C;Genetics:
A;Gene: cyb
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex
with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;1-307/Domain: cytochrome b homology (fragment) <CBH>
F;1-98/Domain: cytochrome b6 homology (fragment) <CB6>
F;4-20/Domain: transmembrane #status predicted <TM1>
F;49-67/Domain: transmembrane #status predicted <TM2>
F;85-101/Domain: transmembrane #status predicted <TM3>
F;146-168/Domain: transmembrane #status predicted <TM4>
F;189-307/Domain: plastocyanin reductase 17K protein homology <17K>
F;197-213/Domain: transmembrane #status predicted <TM5>
P;256-272/Domain: transmembrane #status predicted <TM6>
F;51,150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;65,164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
| | | | |
Db 256 LGGVLA 261

RESULT 352
S22919
ubiquinol-cytochrome-c reductase [EC 1.10.2.2] cytochrome b - creamy-breasted canastero
C;Species: Mitochondrion Asthenes dorbignyi (creamy-breasted canastero)
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 03-Jun-2002
C;Accession: S22919
R;Edwards, S.V.; Arcander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A>Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A;Reference number: S22919; MUID:91288587; PMID:1676522
A;Accession: S22919
A;Status: translation not shown
A:Molecule type: DNA
A;Residues: 1-308 <EDW>
C;Cross-references: UNIPARC:UPI0000128B22; EMBL:X60946; NID:g13749; PIDN:CAA43281.1; PID:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;1-307/Domain: cytochrome b homology (fragment) <CBH>

F;1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F;4-20/Domain: transmembrane #status predicted <TM1>
F;49-67/Domain: transmembrane #status predicted <TM2>
F;85-101/Domain: transmembrane #status predicted <TM3>
F;146-168/Domain: transmembrane #status predicted <TM4>
F;189-307/Domain: plastocyanin reductase 17K protein homology <17K>
F;197-213/Domain: transmembrane #status predicted <TM5>
F;256-272/Domain: transmembrane #status predicted <TM6>
F;51.150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted:
F;65.164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted:

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 256 LGGVLA 261

RESULT 353
S22921
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - white-cheeked cotinga mite
C;Species: mitochondrion Ampelion stressmanni (white-cheeked cotinga)
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C;Accession: S22921
R;Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A;Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A;Reference number: S22919; PMID:91288587; PMID:1676522
A;Accession: S22921
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <EDW>
A;Cross-references: UNIPROT:P29633; UNIPARC:UPI000012890A; EMBL:X60947; NID:gl2746; PID:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion,
F;1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F;4-20/Domain: transmembrane #status predicted <TM1>
F;49-67/Domain: transmembrane #status predicted <TM2>
F;85-101/Domain: transmembrane #status predicted <TM3>
F;146-168/Domain: transmembrane #status predicted <TM4>
F;189-307/Domain: plastocyanin reductase 17K protein homology <17K>
F;197-213/Domain: transmembrane #status predicted <TM5>
F;256-272/Domain: transmembrane #status predicted <TM6>
F;51.150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted:
F;65.164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted:

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 256 LGGVLA 261

RESULT 354
S22925
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Australian magpie mitochondrion
C;Species: mitochondrion Gymnorhina tibicen (Australian magpie)
C;Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C;Accession: S22925
R;Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A;Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A;Reference number: S22919; PMID:91288587; PMID:1676522
A;Accession: S22925
A;Status: translation not shown

QY 17 LGGVLA 22
Db 256 LGGVLA 261

RESULT 358
S22930
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - hooded pitta mitochondrion
C:Species: mitochondrion Pitta sordida (hooded pitta)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C:Accession: S22930
R:Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A:Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A:Reference number: S22919; MUID:91288587; PMID:1676522
A:Accession: S22930
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-308 <EDW>
A:Cross-references: UNIPROT:P29639; UNIPARC:UPI0000128A97; EMBL:X60948; NID:g13382; PIDN
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:1-307/Domain: cytochrome b homology (fragment) <CBH>
F:1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F:4-20/Domain: transmembrane #status predicted <TM1>
F:49-67/Domain: transmembrane #status predicted <TM2>
F:85-101/Domain: transmembrane #status predicted <TM3>
F:146-168/Domain: transmembrane #status predicted <TM4>
F:189-307/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:197-213/Domain: transmembrane #status predicted <TM5>
F:256-272/Domain: transmembrane #status predicted <TM6>
F:51,150/Domain: transmembrane #status predicted <TM6>
F:51,150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:65,164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 256 LGGVLA 261

RESULT 359
S22920
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - MacGregor's bowerbird mito
C:Species: mitochondrion Amblyornis macgregoriae (MacGregor's bowerbird)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C:Accession: S22920
R:Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A:Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A:Reference number: S22919; MUID:91288587; PMID:1676522
A:Accession: S22920
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-308 <EDW>
A:Cross-references: UNIPROT:P29632; UNIPARC:UPI0000128905; EMBL:X60940; NID:g12683; PIDN
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:1-307/Domain: cytochrome b homology (fragment) <CBH>
F:1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F:4-20/Domain: transmembrane #status predicted <TM1>
F:49-67/Domain: transmembrane #status predicted <TM2>
F:85-101/Domain: transmembrane #status predicted <TM3>
F:146-168/Domain: transmembrane #status predicted <TM4>

F:189-307/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:197-213/Domain: transmembrane #status predicted <TM5>
F:256-272/Domain: transmembrane #status predicted <TM6>
F:51,150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:65,164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 256 LGGVLA 261

RESULT 360
S22932
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Andean tapaculo mitochondri
C:Species: mitochondrion Scytalopus magellanicus (Andean tapaculo)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C:Accession: S22932
R:Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A:Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A:Reference number: S22919; MUID:91288587; PMID:1676522
A:Accession: S22932
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-308 <EDW>
A:Cross-references: UNIPROT:P29641; UNIPARC:UPI0000128AE2; EMBL:X60945; NID:g13631; PIDN
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:1-307/Domain: cytochrome b homology (fragment) <CBH>
F:1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F:4-20/Domain: transmembrane #status predicted <TM1>
F:49-67/Domain: transmembrane #status predicted <TM2>
F:85-101/Domain: transmembrane #status predicted <TM3>
F:146-168/Domain: transmembrane #status predicted <TM4>
F:189-307/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:197-213/Domain: transmembrane #status predicted <TM5>
F:256-272/Domain: transmembrane #status predicted <TM6>
F:51,150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:65,164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 256 LGGVLA 261

RESULT 361
S22928
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rufous babbler mitochondri
C:Species: mitochondrion Pomatostomus isidorei (rufous babbler)
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
C:Accession: S22928; I33285; S21641; S21642; S70407
R:Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
A:Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A:Reference number: S22919; MUID:91288587; PMID:1676522
A:Accession: S22928
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-308 <EDW>
A:Cross-references: UNIPROT:P16363; UNIPARC:UPI0000128AA7; EMBL:X60938; NID:g13345; PIDN
R:Kocher, T.D.; Thomas, W.K.; Meyer, A.; Edwards, S.V.; Paeaebo, S.; Villablanca, P.X.;
Proc. Natl. Acad. Sci. U.S.A. 86, 6196-6200, 1989

A;Title: Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing
A;Reference number: A33285; MUID:89345630; PMID:2762322
A;Accession: J33285
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 15,'A',17-89,'T',91-94 <KOC>
R;Edwards, S.V.; Wilson, A.C.
A;Cross-references: UNIPARC:UPI000016D5EB; GB:M25689; NID:g343670; PIDN:AAA32140.1; PID:
submitted to the EMBL Data Library, October 1990
A;Description: Phylogenetically informative length polymorphism and sequence variability
A;Reference number: S21611
A;Accession: S21641
A;Molecule type: DNA
A;Residues: 5-6,'I',8-98 <ED3>
A;Cross-references: UNIPARC:UPI000009950A; EMBL:X54887; NID:gl3341; PIDN:CAA38660.1; PID
A;Experimental source: strain RU692
A;Accession: S21642
A;Molecule type: DNA
A;Residues: 5-6,'I',8-98 <EDP>
A;Cross-references: UNIPARC:UPI000009950A; EMBL:X54886; NID:gl3347; PIDN:CAA38659.1; PID
A;Experimental source: strain RU796
R;Edwards, S.V.; Wilson, A.C.
Genetics 126, 695-711, 1990
A;Title: Phylogenetically informative length polymorphism and sequence variability in mi
A;Reference number: S70404; MUID:91065505; PMID:1979038
A;Accession: S70407
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 5-6,'I',8-98 <ED2>
A;Cross-references: UNIPARC:UPI000009950A; EMBL:X54887; NID:gl3341; PIDN:CAA38660.1; PID
A;Experimental source: isolate RU692; isolate RU796
C;Genetics:
A;Genome: mitochondrion
A;Superfamily: SGC1
C;Keywords: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
in
F;1-307/Domain: cytochrome b homology (fragment) <CBH>
F;1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F;4-20/Domain: transmembrane #status predicted <TM1>
F;49-67/Domain: transmembrane #status predicted <TM2>
F;85-101/Domain: transmembrane #status predicted <TM3>
F;146-168/Domain: transmembrane #status predicted <TM4>
F;189-307/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;197-213/Domain: transmembrane #status predicted <TM5>
F;256-272/Domain: transmembrane #status predicted <TM6>
F;51.150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;65.164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 256 LGGVLA 261

RESULT 362
S22926
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - chestnut-crowned babbler m
C;Species: mitochondrion Pomatostomus ruficeps (chestnut-crowned babbler)
C;Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text_change 09-Jul-2004
C;Accession: S22926; F33285; S70409; S70406
R;Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. Biol. Sci. 243, 99-107, 1991
A;Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A;Reference number: S22919; MUID:91288587; PMID:1676522
A;Accession: S22926
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <EDW>
A;Cross-references: UNIPARC:UPI0000128AA9; EMBL:X60937; NID:gl3316; PIDN

R;Kocher, T.D.; Thomas, W.K.; Meyer, A.; Edwards, S.V.; Paeaebo, S.; Villablanca, F.X.;
Proc. Natl. Acad. Sci. U.S.A. 86, 6196-6200, 1989
A;Title: Dynamics of mitochondrial DNA evolution in animals: amplification and sequencin
A;Reference number: A33285; MUID:89345630; PMID:2762322
A;Accession: F33285
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 15,'A',17-94 <KOC>
A;Cross-references: UNIPARC:UPI000016D60C; GB:M25686; NID:g343664; PIDN:AAA32137.1; PID:
R;Edwards, S.V.; Wilson, A.C.
Genetics 126, 695-711, 1990
A;Title: Phylogenetically informative length polymorphism and sequence variability in mi
A;Reference number: S70404; MUID:91065505; PMID:1979038
A;Accession: S70409
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 5-34,'T',36-98 <EDF>
A;Cross-references: UNIPARC:UPI0000095855; EMBL:X54911; NID:gl3491; PIDN:CAA38683.1; PID
A;Experimental source: individual 19B
A;Accession: S70406
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 5-98 <EDA>
A;Cross-references: UNIPARC:UPI000016D60B; EMBL:X54910; EMBL:X54912; NID:gl3376; PIDN:CA
A;Experimental source: individual 18F
C;Genetics:
A;Gene: cyb
A;Genome: mitochondrion
A;Superfamily: SGC1
C;Keywords: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;1-307/Domain: cytochrome b homology (fragment) <CBH>
F;1-177/Domain: cytochrome b6 homology (fragment) <CB6>
F;4-20/Domain: transmembrane #status predicted <TM1>
F;49-67/Domain: transmembrane #status predicted <TM2>
F;85-101/Domain: transmembrane #status predicted <TM3>
F;146-168/Domain: transmembrane #status predicted <TM4>
F;189-307/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;197-213/Domain: transmembrane #status predicted <TM5>
F;256-272/Domain: transmembrane #status predicted <TM6>
F;51.150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;65.164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 256 LGGVLA 261

RESULT 363
S22922
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - hermit thrush mitochondrion
C;Species: mitochondrion Catharus guttatus (hermit thrush)
C;Date: 29-Jan-1993 #sequence revision 29-Jan-1993 #text_change 09-Jul-2004
C;Accession: S22922
R;Edwards, S.V.; Arctander, P.; Wilson, A.C.
Proc. R. Soc. Lond. Biol. Sci. 243, 99-107, 1991
A;Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
A;Reference number: S22919; MUID:91288587; PMID:1676522
A;Accession: S22922
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <EDW>
A;Cross-references: UNIPROT:P29634; UNIPARC:UPI0000128970; EMBL:X60939; NID:gl2867; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Superfamily: SGC1
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;1-307/Domain: cytochrome b homology (fragment) <CBH>

F:1-177/Domain: cytochrome b6 homology (fragment) <CB6>
 F:4-20/Domain: transmembrane #status predicted <TM1>
 F:49-67/Domain: transmembrane #status predicted <TM2>
 F:95-101/Domain: transmembrane #status predicted <TM3>
 F:146-168/Domain: transmembrane #status predicted <TM4>
 F:189-307/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
 F:197-213/Domain: transmembrane #status predicted <TM5>
 F:256-272/Domain: transmembrane #status predicted <TM6>
 F:65,164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted
 F:65,164/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 256 LGGVLA 261

RESULT 364
 S22923
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Andean flicker mitochondrion
 C:Species: mitochondrion Colaptes rupicola (Andean flicker)
 C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 09-Jul-2004
 C:Accession: S22923
 R:Edwards, S.V.; Arctander, P.; Wilson, A.C.
 Proc. R. Soc. Lond. B Biol. Sci. 243, 99-107, 1991
 A:Title: Mitochondrial resolution of a deep branch in the genealogical tree for perching
 A:Reference number: S22919; MUID:91288587; PMID:1676522
 A:Accession: S22923
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-308 <EDW>
 A:Cross-references: UNIPROT:P29635; UNIPARC:UPI0000128986; EMBL:X60949; NID:g12892; PIDN
 A:Genome: mitochondrion
 A:Genetic code: SGCI
 A:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
 A:Keywords: Chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:1-307/Domain: cytochrome b homology (fragment) <CB>
 F:1-177/Domain: cytochrome b6 homology (fragment) <CB6>
 F:4-20/Domain: transmembrane #status predicted <TM1>
 F:49-67/Domain: transmembrane #status predicted <TM2>
 F:85-101/Domain: transmembrane #status predicted <TM3>
 F:146-168/Domain: transmembrane #status predicted <TM4>
 F:189-307/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
 F:197-213/Domain: transmembrane #status predicted <TM5>
 F:256-272/Domain: transmembrane #status predicted <TM6>
 F:65,164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted
 F:65,164/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 256 LGGVLA 261

RESULT 365
 D95932
 probable sugar uptake ABC transporter permease protein Smb21220 [imported] - Sinorhizobium
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C:Accession: D95932
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: D95932
 A:Status: preliminary

A:Molecule type: DNA
 A:Residues: 1-308 <KUR>
 A:Cross-references: UNIPROT:Q92VH9; UNIPARC:UPI00000CB62D; GB:AL591985; PIDN:CAC49124.1,
 R:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, P.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.,
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Leilaure,
 hehault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Smb21220
 A:Genome: plasmid
 C:Superfamily: inner membrane protein ugpa

Query Match 5.1%; Score 6; DB 2; Length 308;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAL 24
 |||||
 Db 101 GVLAL 106

RESULT 366
 AB0343
 probable phosphatidate cytidyltransferase (SC 2.7.7.41) [imported] - Yersinia pestis
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AB0343
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AB0343
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-309 <KUR>
 A:Cross-references: UNIPROT:Q8ZCV9; UNIPARC:UPI00000DCB15; GB:AL590842; PIDN:CAC93049.1,
 A:Gene: YPO2816
 C:Superfamily: phosphatidate cytidyltransferase
 C:Keywords: nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 309;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||
 Db 220 LLGGVL 225

RESULT 367
 A71374
 probable hypothetical protein - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C:Accession: A71374
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: A71374
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

A;Residues: 1-310 <COL>
A;Cross-references: UNIPROT:Q5G345; UNIPARC:UPI000000D3B18; GB:AE001189; GB:AE000520; NID
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0042
C;Superfamily: Treponema pallidum hypothetical protein TP0042

Query Match 5.1%; Score 6; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLAGVVL 21
Db 164 LLAGVVL 169
|||||

RESULT 368
A95932
probable sugar kinase protein (EC 2.7.1.-) [imported] - Sinorhizobium meliloti (strain 1
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A95932
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, P.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 96, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-Kb pSymB megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: A95932
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-311 <KUR>
A;Cross-references: UNIPROT:Q92VI2; UNIPARC:UPI000000CB62A; GB:AL591985; PIDN:CAC49121.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMB21217
A;Genome: plasmid
C;Keywords: phosphotransferase

Query Match 5.1%; Score 6; DB 2; Length 311;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 267 LAALAA 272
|||||

RESULT 369
C70475
cytochrome c biogenesis protein - Aquifex aeolicus
C;Species: Aquifex aeolicus
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C;Accession: C70475
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196866; PMID:9537320
A;Accession: C70475
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-312 <AQF>
A;Cross-references: UNIPROT:Q67831; UNIPARC:UPI000005679B; GB:AE000769; NID:g2984262; PI
A;Experimental source: strain VF5
C;Genetics:

A;Gene: hemX2
C;Superfamily: cytochrome c-type synthesis protein

Query Match 5.1%; Score 6; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLAGVVL 21
Db 20 LLAGVVL 25
|||||

RESULT 370

T35111

probable tRNA delta(2)-isopentenylpyrophosphate transferase - Streptomyces coelicolor
C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T35111

R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, March 1998

A;Reference number: Z21568

A;Accession: T35111

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-312 <SEE>

A;Cross-references: UNIPROT:O69967; UNIPARC:UPI000012F0D3; EMBL:AL022268; PIDN:CAA18328.

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: miaA; SC08DB:SC4H2.12

C;Superfamily: delta(2)-isopentenylpyrophosphate transferase

Query Match 5.1%; Score 6; DB 2; Length 312;

Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 247 VLAALA 252
|||||

RESULT 371

E64072

spermidine/putrescine-binding protein 2 precursor HI0498 - Haemophilus influenzae (strai
C;Species: Haemophilus influenzae

C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 05-Oct-2004

C;Accession: E64072

R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, .

A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A;Reference number: A64000; MUID:95350630; PMID:7542800

A;Accession: E64072

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-312 <TIGR>

A;Cross-references: UNIPARC:UPI00001685A6; GB:U32731; GB:L42023; NID:g3212193; PIDN:AAC2

C;Genetics:

A;Gene: potD; potD-A

A;Start codon: GTG

C;Function:

A;Description: primary receptor of the polyamine transport system which regulates the po

C;Superfamily: spermidine/putrescine-binding protein

C;Keywords: periplasmic space; transport protein

Query Match 5.1%; Score 6; DB 2; Length 312;

Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PKGKVL 87
Db 124 PKGKVL 129
|||||

RESULT 372

C87562
conserved hypothetical protein CC2524 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: C87562
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.B.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87562
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <STO>
A:Cross-references: UNIPROT:Q9A5C6; UNIPARC:UPI00000C7768; GB:AE005673; NID:gl3424083; F
C:Genetics:
A:Gene: CC2524

Query Match 5.1%; Score 6; DB 2; Length 312;

Best Local Similarity 100.0%; Pred. No. 4.7e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25

|||||

Db 228 VLAALA 233

RESULT 373

G97382
probable ATP-binding component of a transport system (AE005444) [imported] - Agrobacteri
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004
C:Accession: G97382
R.Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97382
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-312 <KUR>
A:Cross-references: UNIPROT:Q8UIU5; UNIPARC:UPI00000D171E; GB:AE007869; PIDN:AAK86016.1;
C:Genetics:
A:Gene: AGR_C_334
A:Map position: circular chromosome

Query Match 5.1%; Score 6; DB 2; Length 312;

Best Local Similarity 100.0%; Pred. No. 4.7e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRAT 94

|||||

Db 304 LLQRAT 309

RESULT 374

AF2600
hypothetical protein Atu0197 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 31-Dec-2004
C:Accession: AF2600
R.Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

C;Superfamily: lysophospholipase L2

Query Match 5.1%; Score 6; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 133 LAALAA 138

RESULT 377

D96028
probable transcription activator of the pca operon, LysR family protein [imported] - Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: D96028
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: D96028
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-313 <KUR>
A;Cross-references: UNIPROT:Q927L9; UNIPARC:UPI00000CB8A7; GB:AL591985; PIDN:CAC49892.1;
A;Experimental source: strain 1021, megaplasmid pSymb
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: pcaQ; SMB20580
A;Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||||
Db 260 GVLAAL 265

RESULT 378

NJBYM1
toxin M1-1 precursor - Saccharomyces cerevisiae killer particle M1
C;Species: Saccharomyces cerevisiae killer particle M1_ScV-M1
C;Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 18-Jun-1999
C;Accession: A01809; A90844; A90984; S72158
R;Boatman, K.A.; Elliott, Q.; Bussey, H.; Burn, V.; Smith, A.; Tipper, D.J.
Cell 36, 741-751, 1984
A;Title: Sequence of the preprotoxin dRNA gene of type I killer yeast: multiple processes
A;Reference number: A90844; MUID:84130185; PMID:6697395
A;Accession: A01809
A;Molecule type: mRNA
A;Residues: 1-316 <BOS>
A;Cross-references: UNIPARC:UPI000012EAAE; EMBL:K02042; NID:g171869; PIDN:AAA34748.1; PI
A;Accession: A90844
A;Molecule type: protein
A;Residues: 1, 'XLXXX', 7-9, 'XX', 12, 'XX', 15, 'XXXX', 20-21, 'XX', 24-25, 45-64, 234-238, 'X', 240-
A;Cross-references: UNIPARC:UPI0000173625; UNIPARC:UPI0000173626; UNIPARC:UPI0000173627
R;Skjipper, N.; Thomas, D.Y.; Lau, P.C.K.
EMBO J. 3, 107-111, 1984
A;Title: Cloning and sequencing of the preprotoxin-coding region of the yeast M1 double-
A;Reference number: A90984; MUID:84158518; PMID:6368221
A;Accession: A90984
A;Molecule type: mRNA

A;Residues: 1-102, 'S', 104-122, 'A', 124-316 <SKI>
A;Cross-references: UNIPARC:UPI0000168BEE; EMBL:X00285; NID:g3663; PIDN:CAA25078.1; PID:
R;Russell, R.J.; Bennett, A.M.; Love, Z.; Baggett, D.M.
submitted to the EMBL Data Library, November 1996
A;Description: Cloning, sequencing and expression of a full-length cDNA copy of the M1 d

A;Reference number: S72158
A;Accession: S72158
A;Molecule type: genomic RNA
A;Residues: 1-316 <RUS>

A;Cross-references: UNIPARC:UPI000013EAAE; GB:U78817; NID:g1699029; PIDN:AAC58005.1; PID:
C;Superfamily: M1-1 toxin
C;Keywords: glycoprotein; transmembrane protein
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-316/Product: protoxin #status predicted <PRX>
F;27-44/Domain: propeptide (delta chain) #status predicted <PRD>
F;45-130/Product: toxin M1-1 alpha chain #status predicted <ACH>
F;72-91/Domain: transmembrane #status predicted <TM1>
F;112-123/Domain: transmembrane #status predicted <TM2>
F;131-233/Domain: propeptide (gamma chain) #status predicted <PRG>
F;234-316/Product: toxin M1-1 beta chain #status experimental <BCH>
F;92-248,95-312,107-239/Disulfide bonds: #status predicted
F;181,203,216/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 316;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 LLQRAT 94
|||||
Db 59 LLQRAT 64

RESULT 379

B83658
phosphoribosyl pyrophosphate synthetase prs [imported] - Bacillus halodurans (strain C-1
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B83658
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-316 <STO>
A;Cross-references: UNIPROT:Q9KGJ5; UNIPARC:UPI000012DFA9; GB:AP001507; GB:BA000004; NID:
A;Experimental source: strain C-125
C;Genetics:
A;Gene: prs
C;Superfamily: ribose-phosphate pyrophosphokinase catalytic chain

Query Match 5.1%; Score 6; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 36 IVGHIE 41
|||||
Db 210 IVGHIE 215

RESULT 380

H95985
probable transcription regulator, araC family protein [imported] - Sinorhizobium meliloti
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: H95985
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: H95985
A;Status: preliminary

A:Molecule type: DNA
 A:Cross-references: UNIPROT:Q92UH9; UNIPARC:UPI00000CB782; GB:AL591985; PIDN:CAC49552.1;
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMB20859
 A:Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 316;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 232 LAALAA 237

RESULT 381

H82785
 dolichol-phosphate mannosyltransferase XF0612 [imported] - Xylella fastidiosa (strain 98)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C:Accession: H82785
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: H82785
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-317 <SIM>
 A:Cross-references: UNIPROT:Q9PPF5; UNIPARC:UPI00000C248A; GB:AE0033906; GB:AE0033849; NID
 A:Experimental source: strain 985c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorry, H.; Pacincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0612

Query Match 5.1%; Score 6; DB 2; Length 317;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 246 LGGVLA 251

RESULT 382

T45397
 FixB [imported] - Mycobacterium leprae
 C:Species: Mycobacterium leprae

C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 31-Dec-2004
 C:Accession: T45397
 R:Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, September 1997
 A:Reference number: Z16918
 A:Accession: T45397
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-318 <PAR>
 A:Cross-references: UNIPROT:Q33096; UNIPARC:UPI000012A248; EMBL:Z99263; PIDN:CAB16419.1
 A:Experimental source: cosmid B637
 C:Genetics:
 A:Note: fixB
 C:Superfamily: electron transfer flavoprotein, alpha subunit

Query Match 5.1%; Score 6; DB 2; Length 318;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
 |||||
 Db 77 VLAALA 82

RESULT 383

T36438
 probable membrane protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T36438
 R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, July 1999
 A:Reference number: Z21598
 A:Accession: T36438
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-318 <SEE>
 A:Cross-references: UNIPROT:Q9XAA0; UNIPARC:UPI00000DB21A; EMBL:AL096837; PIDN:CAB48902.
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCF43A.15

Query Match 5.1%; Score 6; DB 2; Length 318;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAA 27
 |||||
 Db 27 AALAA 32

RESULT 384

B75495
 conserved hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: B75495
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: B75495
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-319 <WHI>
 A:Cross-references: UNIPROT:Q9RWNO; UNIPARC:UPI00000C17BE; GB:AE001921; GB:AE000513; NID
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0636
 A:Map position: 1
 C:Superfamily: methyltransferase, YraL type

Query Match 5.1%; Score 6; DB 2; Length 319;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 |||||
 Db 171 LAALAA 176

RESULT 385

151030
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - boatzin mitochondrion (fra
 C;Species: mitochondrion Opisthocomus hoazin (hoatzin)
 C;Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 09-Jul-2004
 C;Accession: 151030
 R;Avise, J.C.; Nelson, W.S.; Sibley, C.G.

Mol. Phylogenet. Evol. 3, 175-184, 1994
 A;Title: Why one-kilobase sequences from mitochondrial DNA fail to solve the Hoatzin phy
 A;Reference number: 150122; MUID:94356264; PMID:8075835

A;Accession: 151030

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-320 <AVI>

A;Cross-references: UNIPROT:Q37124; UNIPARC:UPI0000092P73; EMBL:U09257; NID:g769841; PID

C;Genetics:

A;Genome: mitochondrion

A;Genetic code: SGC1

C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;

F;1-307/Domain: cytochrome b homology (fragment) <CB6>

F;1-177/Domain: cytochrome b6 homology (fragment) <CB6>

F;189-307/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F;51.150/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F;65.164/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 256 LGGVLA 261

RESULT 386

S69547
 transcription initiation factor sigma 2 - Synechococcus sp. (strain PCC 7942)

A;Variety: PCC 7942

C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004

C;Accession: S69547; JN0332

R;Tsinozemas, N.F.; Ishiura, M.; Kondo, T.; Andersson, C.R.; Tanaka, K.; Takahashi, H.;

EMBO J. 15, 2488-2495, 1996

A;Title: A sigma factor that modifies the circadian expression of a subset of genes in c

A;Reference number: S69547; MUID:96221309; PMID:8665856

A;Accession: S69547

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: DNA

A;Residues: 1-320 <TST>

A;Cross-references: UNIPROT:Q55350; UNIPARC:UPI00000BF060; EMBL:D78583; NID:g2385505; PI

R;Tanaka, K.; Masuda, S.; Takahashi, H.

Biosci. Biotechnol. Biochem. 56, 1113-1117, 1992

A;Title: Multiple rpoD-related genes of cyanobacteria.

A;Reference number: JN0331; MUID:93005065; PMID:1368828

A;Accession: JN0332

A;Molecule type: DNA

A;Residues: 112-159 <TAN>

A;Cross-references: UNIPARC:UPI0000175772; GB:S44854; NID:g256268; PIDN:AAC60415.1; PID:

C;Genetics:

A;Gene: rpoD2; rpoD

C;Superfamily: transcription initiation factor sigma katF; transcription initiation fact

C;Keywords: DNA binding; sigma factor; transcription initiation

F;91-316/Domain: transcription initiation factor sigma katP homology <KTF>

Query Match 5.1%; Score 6; DB 2; Length 320;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 95 QQQAVI 100
 |||||
 Db 263 QQQAVI 268

RESULT 387

G98067
 sucrose regulon regulatory protein [imported] - Streptococcus pneumoniae (strain R6)

C;Species: Streptococcus pneumoniae

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C;Accession: G98067

R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E

e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Mateushima, P.; McAhren, S.; M

Y, P.; Sun, P.M.; Winkler, M.E.

J. Bacteriol. 183, 5709-5717, 2001

A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;

A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.

A;Reference number: A97872; MUID:21429245; PMID:11544234

A;Accession: G98067

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-321 <KUR>

A;Cross-references: UNIPROT:Q97PB5; UNIPARC:UPI0000051975; GB:AE007317;

C;Genetics:

A;Gene: scrR

C;Superfamily: lac repressor

Query Match 5.1%; Score 6; DB 2; Length 321;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
 |||||
 Db 160 GGVLA 165

RESULT 388

A95201
 sucrose operon repressor [imported] - Streptococcus pneumoniae (strain TIGR4)

C;Species: Streptococcus pneumoniae

C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004

C;Accession: A95201

R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid

on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,

nson, T.; Hickey, E.K.; Holt, I.E.

Science 293, 498-506, 2001

A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,

A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.

A;Reference number: A95000; MUID:21357209; PMID:11463916

A;Accession: A95201

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-321 <KUR>

A;Cross-references: UNIPROT:Q97PB5; UNIPARC:UPI0000051975; GB:AE005672;

A;Experimental source: strain TIGR4

C;Genetics:

A;Gene: SP1725

C;Superfamily: lac repressor

Query Match 5.1%; Score 6; DB 2; Length 321;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
 |||||
 Db 160 GGVLA 165

RESULT 389

G82752
cobalt-zinc-cadmium resistance protein XF0866 [imported] - Xylella fastidiosa (strain 9a)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82752
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: G82752
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <SIM>
A:Cross-references: UNIPROT:O9PFI2; UNIPARC:UPI00000C2550; GB:AE003849; NID:
R:Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Pacincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.B.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tauhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0866
C:Superfamily: zinc transporter Znt-2

Query Match 5.1%; Score 6; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 TWVLG 18
|||||
DB 218 TWVLG 223

RESULT 390

T20272
hypothetical protein C56G7.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20272
R:Percy, C.
submitted to the EMBL Data Library, November 1994
A:Reference number: Z19245
A:Accession: T20272
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-322 <WIL>
A:Cross-references: UNIPROT:O09292; UNIPARC:UPI000013BP30; EMBL:Z46793; PIDN:CAA86771.1;
A:Experimental source: clone C56G7
C:Genetics:
A:Gene: CESP:C56G7.3
A:Map position: 3
A:Introns: 26/3; 225/3; 251/2; 294/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C56G7.3

Query Match 5.1%; Score 6; DB 2; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLQ 91
|||||
DB 195 VLGLQ 200

RESULT 391

A72508
probable cobalamin biosynthesis protein APE2039 - Aeropyrum pernix (strain Kl)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: A72508
R:Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: A72508
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KAW>
A:Cross-references: UNIPROT:Q9YAA0; UNIPARC:UPI000005E1B5; DBJ:AF000063; NID:G5105654;
A:Experimental source: strain Kl
C:Genetics:
A:Gene: APE2039
C:Superfamily: cobalamin biosynthesis protein D

Query Match 5.1%; Score 6; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 310 VLAALA 315

RESULT 392

F75631
iron ABC transporter, permease protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: F75631
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75631
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-324 <WHI>
A:Cross-references: UNIPROT:Q9RZK4; UNIPARC:UPI00000D3B63; GB:AE001826; NID:G6460827; P
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRB0123
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MP1
C:Superfamily: ferrichrome ABC transporter

Query Match 5.1%; Score 6; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAA 27
|||||
DB 125 AALAA 130

RESULT 393

T31474
hypothetical protein Y62P5A.1a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31474; T31475
R:Williams, L.

submitted to the EMBL Data Library, October 1999

A;Reference number: Z21038
A;Accession: T31474
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-325 <WIL>
A;Cross-references: UNIPROT:Q9ULW3; UNIPARC:UPI0000075881; EMBL:AL110499; PIDN: CAB57912.
A;Experimental source: clone Y62F5A
A;Accession: T31475
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-293 'KFOQ' <W12>
A;Cross-references: UNIPARC:UPI0000074803; EMBL:AL110499; PIDN: CAB57913.1; CESP: Y62F5A.1
A;Experimental source: clone Y62F5A
C;Genetics:
A;Gene: CESP: Y62F5A.1a; CESP: Y62F5A.1b
A;Introns: 110/3; 188/3; 294/1

Query Match 5.1%; Score 6; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 90 LQRATQ 95
|||||
Db 109 LQRATQ 114

RESULT 394

G96718
unknown protein, 54453-53476 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G96718
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Holt, D.; Chang, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.; Jones, R.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-A.; Li, J.H.; Li, Y.; Liu, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID: 21016719; PMID: 11130712
A;Accession: G96718
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-325 <STO>
A;Cross-references: UNIPROT:Q9C9L2; UNIPARC:UPI000009C948; GB:AE005173; NID:G6665547; PID:G6665547
C;Genetics:
A;Gene: T6C23.11
A;Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||||
Db 282 GVLAAL 287

RESULT 395

AG1688
protein-tyrosine/serine phosphatase homolog lin2049 [imported] - Listeria innocua (strain ATCC 35061)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AG1688
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A;Title: Comparative genomics of *Listeria species*.
A;Reference number: AB1077; MUID: 21537279; PMID: 11679669
A;Accession: AG1688
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <GLA>
A;Cross-references: UNIPROT:Q92A73; UNIPARC:UPI000000CC733; GB:AL592022; PIDN: CAC97279.1;
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2049

Query Match 5.1%; Score 6; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||||
Db 223 GVLAAL 228

RESULT 396

T44255
probable amino acid oxidase flavoprotein [imported] - Rhizobium etli plasmid b
C;Species: Rhizobium etli
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T44255
R;Miranda-Rios, J.; Moreza, C.; Taboada, H.; Davalos, A.; Encarnacion, S.; Mora, J.; Sob J. Bacteriol. 179, 6887-6893, 1997
A;Title: Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiotic
A;Reference number: 222737; MUID: 98037482; PMID: 9371431
A;Accession: T44255
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-327 <MIR>
A;Cross-references: UNIPROT:O34292; UNIPARC:UPI0000136E97; EMBL:AF004408; NID:G2627325;
A;Experimental source: strain CE3
C;Genetics:
A;Gene: thio
A;Genome: plasmid b
C;Superfamily: D-amino-acid oxidase

Query Match 5.1%; Score 6; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 143 LAALAA 148

RESULT 397

D71491
hypothetical protein CT627 - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C;Species: Chlamydia trachomatis
C;Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 05-Oct-2004
C;Accession: D71491
R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A;Reference number: A71570; MUID: 99000809; PMID: 9784136
A;Accession: D71491
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-327 <ARN>
A;Cross-references: UNIPROT:O84632; UNIPARC:UPI0000139B4A; GB:AE001333; NID:G2627325;
A;Experimental source: serotype D, strain UW-3/Cx
C;Genetics:
A;Gene: CT627
C;Superfamily: uncharacterized conserved protein

Query Match 5.1%; Score 6; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EBCSOA 68
|||||

Db 284 EBCSOA 289

RESULT 398

T30275

hypothetical protein - *Synechococcus* sp. (PCC 7942)

C:Species: *Synechococcus* sp.

A:Variety: PCC 7942

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C:Accession: T30275

R:Phung, L.T.; Haselkorn, R.

submitted to the EMBL Data Library, May 1996

A:Description: Genes encoding biotin carboxyl carrier protein and elongation factor P fr

A:Reference number: Z20804

A:Accession: T30275

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-327 <PHU>

A:Cross-references: UNIPROT:Q54757; UNIPARC:UPI0000136E4D; EMBL:U59235; NID:g1399825; PI

C:Superfamily: thiamine monophosphate kinase

Query Match 5.1%; Score 6; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EQAQVI 78
|||||

Db 290 EQAQVI 295

RESULT 399

DB1650

conserved hypothetical protein TC0916 [imported] - *Chlamydia muridarum* (strain Nigg)

C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 05-Oct-2004

C:Accession: DB1650

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: DB1650

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-328 <TET>

A:Cross-references: UNIPROT:Q9PJB6; UNIPARC:UPI0000057AB8; GB:AE002358; GB:AE002160; NID

A:Experimental source: strain Nigg (MoPn)

C:Genetics:

A:Gene: TC0916

C:Superfamily: uncharacterized conserved protein

Query Match 5.1%; Score 6; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EBCSOA 68
|||||

Db 284 EBCSOA 289

RESULT 400

E82644

sugar-phosphate dehydrogenase XF1724 [imported] - *Xylella fastidiosa* (strain 9a5c)

C:Species: *Xylella fastidiosa*

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 05-Oct-2004

C:Accession: E82644

R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: E82644

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-329 <SIM>

A:Cross-references: UNIPROT:Q9PCQ4; UNIPARC:UPI00000C27BD; GB:AE003849; NID

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, I; as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranee, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, J
A:Authors: Martins, E.M.F.; Matukuma, A.V.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, A.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;
A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1724

C:Superfamily: Aldehyde reductase

Query Match 5.1%; Score 6; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||||

Db 223 LLGGVL 228

RESULT 401

H70744

hypothetical protein RV0493c - *Mycobacterium tuberculosis* (strain H37RV)

C:Species: *Mycobacterium tuberculosis*

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004

C:Accession: H70744

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, K.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: H70744

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-329 <COL>

A:Cross-references: UNIPROT:Q11158; UNIPARC:UPI00001398A2; GB:Z77162; GB:AL123456; NID:9

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0493c

Query Match 5.1%; Score 6; DB 2; Length 329;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIEGG 44
|||||

Db 295 HIEGG 300

RESULT 402

T38355

zinc finger protein - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: T38355

R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1995

A:Reference number: Z21787
A:Accession: T38355
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-329 <MUR>
A:Cross-references: UNIPROT:O13974; UNIPARC:UPI000006AGB7; EMBL:Z98601; PIDN:CAB11271.1;
A:Experimental source: strain 972h-, cosmid C24C9
C:Genetics:
A:Gene: SPDB:SPAC24C9.14
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 329;

Best Local Similarity 100.0%; Pred. No. 4.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 21 LAALAA 26

Db 250 LAALAA 255

RESULT 403

A71365

probable Lambda CII stability-governing protein (hflC) - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004

C:Accession: A71365

R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998

A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:9832770; PMID:9665876

A:Accession: A71365

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-331 <COL>

A:Cross-references: UNIPROT:O83152; UNIPARC:UPI000012C68B; GB:AE001196; GB:AE000520; NID

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0114

Query Match 5.1%; Score 6; DB 2; Length 331;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20

Db 24 VLLGGV 29

RESULT 404

AF3267

peptidyl-prolyl cis-trans isomerase [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004

C:Accession: AF3267

R:DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AB3252; PMID:11756688

A:Accession: AF3267

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-331 <MUR>

A:Cross-references: UNIPROT:O8YJG0; UNIPARC:UPI0000057B68; GB:AE008917; PIDN:AAL51305.1;

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0123

A:Map position: I

Query Match 5.1%; Score 6; DB 2; Length 331;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

Db 98 LAALAA 103

RESULT 405

AI1355

low-affinity inorganic phosphate transporter homolog lmo2249 [imported] - Listeria monoc

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AI1355

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AI1355

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-332 <GLA>

A:Cross-references: UNIPROT:O8Y528; UNIPARC:UPI000005505B; GB:NC_003210; PIDN:CAD00327.1;

A:Experimental source: strain BGD-e

C:Genetics:

A:Gene: lmo2249

C:Superfamily: probable sodium-dependent phosphate transporter MTH1885

Query Match 5.1%; Score 6; DB 2; Length 332;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

Db 12 VLAALA 17

RESULT 406

AB1726

low-affinity inorganic phosphate transporter homolog lin2351 [imported] - Listeria innoc

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C:Accession: AB1726

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1726

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-332 <GLA>

A:Cross-references: UNIPROT:Q929C6; UNIPARC:UPI000000CC832; GB:AL592022; PIDN:CAC97578.1;

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin2351

C:Superfamily: probable sodium-dependent phosphate transporter MTH1885

Query Match 5.1%; Score 6; DB 2; Length 332;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

Db 12 VLAALA 17


```
RESULT 407
B64085
GlpX protein - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 09-Jul-2004
C:Accession: B64085
R:Plieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.;
Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.;
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Frazer, C.M.; Smith, H.O.; Venter,
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: B64085
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-333 <TIGR>
A:Cross-references: UNIPROT:P44811; UNIPARC:UPI000012B7B9; GB:U32749; GB:U42023; NID:G32
C:Genetics:
A:Gene: glpX
C:Superfamily: fructose-1,6-bisphosphatase, GlpX type

Query Match 5.1%; Score 6; DB 2; Length 333;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAY 27
Db 16 AALAAAY 21

RESULT 408
S54438
hemin permease [validated] - Yersinia enterocolitica
C:Species: Yersinia enterocolitica
C>Date: 06-Sep-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C:Accession: S54438
R:Stojiljkovic, I.; Hantke, K.
Mol. Microbiol. 13, 719-732, 1994
A:Title: Transport of haemin across the cytoplasmic membrane through a haemin-specific P
A:Reference number: S54436; MUID:95089707; PMID:7997183
A:Accession: S54438
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-334 <STO>
A:Cross-references: UNIPROT:P74980; UNIPARC:UPI0000178002; EMBL:X77867
C:Genetics:
A:Gene: hemu
C:Function:
A:Description: component of the periplasmic binding-protein-dependent transport system
C:Superfamily: vitamin B12 transport protein btuC
C:Keywords: iron transport; transmembrane protein

Query Match 5.1%; Score 6; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGLLQ 91
Db 206 VLGLLQ 211

RESULT 409
TS0601
endo-1,4-beta-xylanase (EC 3.2.1.8) B, secreted [imported] - Streptomyces coelicolor
N:Alternate names: xylanase B
C:Species: Streptomyces coelicolor
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: TS0601
R:Redenbach, M.; Kleser, H.M.; Denapaite, D.; Eichner, A.; Cullum, J.; Kinaishi, H.; Hopw
Mol. Microbiol. 21, 77-96, 1996
A:Title: A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb S
```

```
A:Reference number: Z20556; MUID:97000351; PMID:8843436
A:Accession: TS0601
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-335 <RED>
A:Cross-references: UNIPROT:Q9RKN6; UNIPARC:UPI000000DC56E; EMBL:AL133220; PIDN:CAB61738
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: xlnB
C:Superfamily: Clostridium endo-1,4-beta-xylanase B; endo-1,4-beta-xylanase homology
C:Keywords: glycosidase; hydrolase

Query Match 5.1%; Score 6; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 27 LAALAA 32

RESULT 410
A40038
MHC class I histocompatibility antigen H-2 M3 alpha chain precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C:Accession: A40038; S09380
R:Wang, C.R.; Loveland, B.E.; Lindahl, K.F.
Cell 66, 335-345, 1991
A:Title: H-2M3 encodes the MHC class I molecule presenting the maternally transmitted al
A:Reference number: A40038; MUID:91309146; PMID:1855254
A:Accession: A40038
A:Molecule type: DNA
A:Residues: 1-335 <MAN>
A:Cross-references: UNIPROT:Q31166; UNIPARC:UPI000008AE1B; GB:M62844; NID:G199407; PIDN
A:Note: the authors translated the codon GTG for residue 160 as Glu
R:Richards, S.; Bucan, M.; Brorson, K.; Kiefer, M.C.; Hunt III, S.W.; Lehrach, H.; Fisci
EMBO J. 8, 3749-3757, 1989
A:Title: Genetic and molecular mapping of the Hmt region of mouse.
A:Reference number: S09380; MUID:90059976; PMID:2573520
A:Accession: S09380
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 207-298 <RIC>
A:Cross-references: UNIPARC:UPI0000176F75
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; heterodimer; transmembrane protein; transplantation antigen
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-335/Product: MHC class I histocompatibility antigen H-2 M3 alpha chain #status pre
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match 5.1%; Score 6; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLGGV 20
Db 316 VLGGV 321

RESULT 411
D70874
probable membrane protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70874
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
```

```
A;Accession: D70874
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-335 <COL>
A;Cross-references: UNIPROT:P71762; UNIPARC:UPI000013AB47; GB:AL021184; GB:AL123456; NID
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv1481

Query Match          5.1%; Score 6; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 321 LAALAA 326
|||||

RESULT 412
H72618
hypothetical protein APE1409 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: H72618
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72618
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-336 <KAW>
A;Cross-references: UNIPROT:Q9YC42; UNIPARC:UPI000005DF33; DDBJ:AP000061; NID:g5104821;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1409
C;Superfamily: Aeropyrum pernix hypothetical protein APE1409

Query Match          5.1%; Score 6; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
Db 86 AALAA 91
|||||

RESULT 413
T23902
hypothetical protein R04D3.8 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T23902
R;Swinburne, J.
A;Submitted to the EMBL Data Library, March 1996
A;Reference number: Z19815
A;Accession: T23902
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-336 <WIL>
A;Cross-references: UNIPROT:Q21716; UNIPARC:UPI000007FC6B; EMBL:Z70212; PIDN:CAA94166.1;
A;Experimental source: clone R04D3
C;Genetics:
A;Gene: CESP:R04D3.8
A;Map position: X
A;Introns: 145/3; 157/3; 222/2; 247/3; 274/1

Query Match          5.1%; Score 6; DB 2; Length 336;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 IEPIVT 105
```

```
Db 294 IEPIVT 299
|||||

RESULT 414
Q0ECH3
Probable dehydrogenase [EC 1.2.1.1]- usg1 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A23792; B33494; I69363; I69364; E65004
R;Arps, P.J.; Marvel, C.C.; Rubin, B.C.; Tolan, D.A.; Penhoet, E.E.; Winkler, M.E.
Nucleic Acids Res. 13, 5297-5315, 1985
A;Title: Structural features of the hist operon of Escherichia coli K-12.
A;Reference number: A93577; MUID:85269644; PMID:2991861
A;Accession: A23792
A;Molecule type: DNA
A;Residues: 1-337 <ARP>
A;Cross-references: UNIPROT:P08390; UNIPARC:UPI0000137E50; GB:X02743; NID:g41716; PIDN:C
A;Experimental source: strain K12
R;Schoenlein, P.V.; Roa, B.B.; Winkler, M.E.
J. Bacteriol. 171, 6084-6092, 1989
A;Title: Divergent transcription of pdxB and homology between the pdxB and serA gene pro
A;Reference number: JVO051; MUID:90036695; PMID:2681152
A;Accession: B33494
A;Molecule type: DNA
A;Residues: 1-27 <SCH>
A;Cross-references: UNIPARC:UPI000016F3AB; GB:M29962; NID:g147122; PIDN:AAA24309.1; PID:
R;Arps, P.J.; Winkler, M.E.
J. Bacteriol. 169, 1061-1070, 1987
A;Title: Structural analysis of the Escherichia coli K-12 hist operon by using a kanamyc
A;Reference number: I54872; MUID:87137258; PMID:3029016
A;Accession: I69363
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-13 <ARP2>
A;Cross-references: UNIPARC:UPI000016F3AF; GB:M15541; NID:g147126; PIDN:AAA24311.1; PID:
A;Accession: I69364
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 329-337 <AR2>
A;Cross-references: UNIPARC:UPI000016F3B0; GB:M15542; NID:g147127; PIDN:AAA24312.1; PID:
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: B65004
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-337 <BLAT>
A;Cross-references: UNIPARC:UPI0000137E50; GB:AE0000320; GB:U00096; NID:g1788647; PIDN:AA
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: usg1
A;Map position: 50 min
C;Superfamily: aspartate-semialdehyde dehydrogenase
C;Keywords: NAD; oxidoreductase

Query Match          5.1%; Score 6; DB 1; Length 337;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 LSVGCV 34
Db 292 LSVGCV 297
|||||

RESULT 415
C91029
probable PTS system enzyme II A component [imported] - Escherichia coli (strain O157:H7;
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C91029
```

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gaeawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic reference number: A99629; PMID:11258796
A:Accession: C91029
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <HAY>
A:Cross-references: UNIPROT:Q8XCR1; UNIPARC:UPI00000D0444; GB:BA000007; PIDN:BA036626.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs3203
C:Superfamily: aspartate-semialdehyde dehydrogenase

Query Match 5.1%; Score 6; DB 2; Length 337;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LSVGCV 34
|||||
Db 292 LSVGCV 297

RESULT 416
D85873
probable PTS system enzyme II A component usg [imported] - *Escherichia coli* (strain O157:H7)
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85873
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; PMID:21074935; PMID:11206551
A:Accession: D85873
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <STO>
A:Cross-references: UNIPROT:Q8XCR1; UNIPARC:UPI0000165893; GB:AE005174; NID:gl2515678; H
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: usg
C:Superfamily: aspartate-semialdehyde dehydrogenase

Query Match 5.1%; Score 6; DB 2; Length 337;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 LSVGCV 34
|||||
Db 292 LSVGCV 297

RESULT 417
A10649
probable regulatory protein STY1297 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhimurium
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A:Note: This species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 17-Mar-2003
C:Accession: A10649
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium
A:Reference number: AB0502; PMID:21534947; PMID:11677608
A:Accession: A10649
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <PAR>
A:Cross-references: UNIPARC:UPI0000059862; GB:AL513382; PIDN:CAD08380.1; PID:gl16502424;

C:Genetics:

A:Gene: STY1297

C:Superfamily: response regulator, Hnr type; response regulator homology

Query Match 5.1%; Score 6; DB 2; Length 337;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24

|||||

Db 212 GVLAAL 217

RESULT 418

H85703

Hnr protein [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004

C:Accession: H85703

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; PMID:21074935; PMID:11206551

A:Accession: H85703

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-337 <STO>

A:Cross-references: UNIPROT:P37055; UNIPARC:UPI000012CA9E; GB:AE005174; NID:gl2514956; H

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: hnr

C:Superfamily: response regulator, Hnr type; response regulator homology

Query Match 5.1%; Score 6; DB 2; Length 337;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24

|||||

Db 212 GVLAAL 217

RESULT 419

A36871

37K regulator response protein homolog - *Escherichia coli* (strain K-12)

N:Alternate names: Hnr protein

C:Species: *Escherichia coli*

C:Date: 21-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004

C:Accession: A36871; S28505; F64870

R:Boal, M.; Kersten, H.

J. Bacteriol. 176, 221-231, 1994

A:Title: Organization and functions of genes in the upstream region of *tyrT* of *Escherichia coli* K-12

A:Reference number: A36871; PMID:94110230; PMID:8282700

A:Accession: A36871

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-337 <BO>

A:Cross-references: UNIPROT:P37055; UNIPARC:UPI000012CA9E; GB:M64675

A:Experimental source: strain K-12

A:Note: sequence extracted from NCBI backbone (NCBI:141999, NCBI:142001)

R:Contreras, A.

submitted to the EMBL Data Library, April 1992

A:Reference number: S28505

A:Accession: S28505

A:Molecule type: DNA

A:Residues: 1-337 <CON>

A:Cross-references: UNIPARC:UPI000012CA9E; EMBL:X66003; NID:g41733; PIDN:CAA46802.1; PII

A:Experimental source: strain K-12

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: F64870

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-337 <BLAT>

A;Cross-references: UNIPARC:UPI000012CA98; GB:AB000222; GB:U00096; MID:gl787486; PIDN:AA

A;Experimental source: strain K-12, substrain MG1655

C;Genetics:

A;Gene: hnr

A;Map position: 28 min

C;Superfamily: response regulator, Hnr type; response regulator homology

C;Keywords: DNA binding; phosphoprotein; transcription regulation

F;10-113/Domain: response regulator homology <RRH>

F;58/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 337;

Best Local Similarity 100.0%; Pred. No. 5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24

|||||

Db 212 GVLAAL 217

RESULT 420

A90846

Hnr protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004

C;Accession: A90846

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: A90846

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-337 <HAY>

A;Cross-references: UNIPROT:P37055; UNIPARC:UPI000012CA9E; GB:BA000007; PIDN:BA035160.1;

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECSA1737

C;Superfamily: response regulator, Hnr type; response regulator homology

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 337;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24

|||||

Db 212 GVLAAL 217

RESULT 421

A10264

probable response regulator YPO2173 [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C;Accession: A10264

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.

dano-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;

Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: A10264

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-338 <KUR>

A;Cross-references: UNIPROT:Q8ZEJ4; UNIPARC:UPI00000DCAAF; GB:AL590842; PIDN:CAC90981.1;

C;Genetics:

A;Gene: YPO2173

C;Superfamily: response regulator, Hnr type; response regulator homology

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 338;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24

|||||

Db 212 GVLAAL 217

RESULT 422

S62596

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Mycena galopoda mitochondr

C;Species: mitochondrion Mycena galopoda

C;Date: 19-Mar-1997 #sequence_revision 28-Aug-1997 #text_change 09-Jul-2004

C;Accession: S62596

R;Kraicz, P.; Haase, U.; Gencic, S.; Flindt, S.; Anke, T.; Brandt, U.; von Jagow, G.

Eur. J. Biochem. 235, 54-63, 1996

A;Title: The molecular basis for the natural resistance of the cytochrome bc(1) complex

A;Reference number: S62595; MUID:96202917; PMID:8631367

A;Accession: S62596

A;Status: nucleic acid sequence not shown

A;Molecule type: nucleic acid

A;Residues: 1-339 <KBA>

A;Cross-references: UNIPROT:Q36443; UNIPARC:UPI0000090B04; EMBL:X87997; NID:9887553; PID

C;Genetics:

A;Gene: cytb

A;Genome: mitochondrion

A;Genetic code: SGC3

A;Introns: 84/3; 227/1

C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

F;1-294/Domain: cytochrome b homology (fragment) <CBH>

F;1-163/Domain: cytochrome b6 homology (fragment) <CB6>

F;33-51/Domain: transmembrane #status predicted <TM1>

F;71-87/Domain: transmembrane #status predicted <TM2>

F;132-154/Domain: transmembrane #status predicted <TM3>

F;175-239/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F;183-194/Domain: transmembrane #status predicted <TM4>

F;242-258/Domain: transmembrane #status predicted <TM5>

F;278-298/Domain: transmembrane #status predicted <TM6>

F;308-324/Domain: transmembrane #status predicted <TM7>

F;35,136/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F;49,150/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 339;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

|||||

Db 142 VLAALA 147

RESULT 423

T34925

ABC transporter integral membrane protein - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T34925

R;Seeger, K.J.; Harris, D.; Parkhill, J.; Barrall, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1998

A;Reference number: Z21562

A;Accession: T34925

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-339 <SEE>

A;Cross-references: UNIPROT:O69943; UNIPARC:UPI00000DAC88; EMBL:AL023862; PIDN:CAAL19626

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SC08DB:SC3F9.03

C;Superfamily: l-arabinose transport system permease araH

Query Match 5.1%; Score 6; DB 2; Length 339;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||||
Db 40 GVLAAL 45

RESULT 424
JN0527
tcpE protein precursor - Vibrio cholerae
C:Species: Vibrio cholerae
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: JN0527
R:Kaufman, M.R.; Shaw, C.E.; Jones, I.D.; Taylor, R.K.
Gene 126, 43-49, 1993
A:Title: Biogenesis and regulation of the Vibrio cholerae toxin-coregulated pilus: Analogs
A:Reference number: JN0527; MUID:93231536; PMID:8097177
A:Accession: JN0527
A:Molecule type: DNA
A:Residues: 1-340 <KAU>
A:Cross-references: UNIPROT:P29487; UNIPARC:UPI0000170594; GB:M93963; NID:g155279; PIDN:
C:Genetics: A:Gene: tcpE
P:1-14/Domain: signal sequence #status predicted <SIG>
F:15-340/Product: tcpE protein #status predicted <MAT>

Query Match 5.1%; Score 6; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVL 55
|||||
Db 71 PDKEVL 76

RESULT 425
JN0525
toxin co-regulated pilus biosynthesis protein E VC0836 [imported] - Vibrio cholerae (str
C:Species: Vibrio cholerae
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: JN0525; S23272; G82275
R:Ogierman, M.A.; Zabih, S.; Mourtzios, L.; Manning, P.A.
Gene 146, 51-60, 1993
A:Title: Genetic organization and sequence of the promoter-distal region of the tcp gene
A:Reference number: JN0521; MUID:93231537; PMID:8097178
A:Accession: JN0525
A:Molecule type: DNA
A:Residues: 1-340 <OGI>
A:Cross-references: UNIPROT:P29487; UNIPARC:UPI0000000A92; EMBL:X64098; NID:g48404; PIDN:
R:Manning, P.A.
submitted to the EMBL Data Library, January 1992
A:Reference number: S23261
A:Accession: S23272
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 15-340 <MAN>
A:Cross-references: UNIPARC:UPI000017AAD5; EMBL:X64098
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82275
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-340 <HEI>
A:Cross-references: UNIPARC:UPI0000000A92; GB:AE004168; GB:AE003852; NID:g9655274; PIDN:
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:

F:313-329/Domain: transmembrane #status predicted <TM8>
F:42,143/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:56,157/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 248 LGGVLA 253
|||||

RESULT 428
F83126
ferric enterobactin transport protein FepG PA4161 [imported] - Pseudomonas aeruginosa (E)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83126
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83126
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <STO>
A:Cross-references: UNIPROT:Q9HWM0; UNIPARC:UPI00000C5C46; GB:AE004832; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: fepG; PA4161

Query Match 5.1%; Score 6; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 133 LAALAA 138
|||||

RESULT 429
T26784
hypothetical protein Y40B1B.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26784
R:Harris, B.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20265
A:Accession: T26784
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-343 <WIL>
A:Cross-references: UNIPROT:Q9XWP5; UNIPARC:UPI0000077EDE; EMBL:AL032636; PIDD:CAA21605.
A:Experimental source: clone Y40B1B
C:Genetics:
A:Gene: CESP:Y40B1B.3
A:Map position: 1
A:Introns: 26/2

Query Match 5.1%; Score 6; DB 2; Length 343;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 AQVIAH 80
Db 118 AQVIAH 123
|||||

RESULT 430

C33830
cation efflux system membrane protein czcC - Alcaligenes eutrophus
C:Species: Alcaligenes eutrophus
C:Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 17-May-2002
C:Accession: C33830
R:Nies, D.H.; Nies, A.; Chu, L.; Silver, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7351-7355, 1989
A:Title: Expression and nucleotide sequence of a plasmid-determined divalent cation effl
A:Reference number: A33830; MUID:90017477; PMID:2678100
A:Accession: C33830
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-345 <NIE>
A:Cross-references: UNIPARC:UPI00001766DE; GB:M26073
C:Superfamily: cyas protein
C:Keywords: membrane protein

Query Match 5.1%; Score 6; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 IELOGK 45
Db 33 IELOGK 38
|||||

RESULT 431
A82348
probable ADP-heptose-LPS heptosyltransferase II VC0223 [imported] - Vibrio cholerae (str
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82348
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; B
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: A82348
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-345 <HEI>
A:Cross-references: UNIPROT:Q9KVC3; UNIPARC:UPI00000C2C2E; GB:AE004112; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 24
Db 240 GVLAAAL 245
|||||

RESULT 432
AD2991
ABC transporter, membrane spanning protein Atu3534 [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AD2991
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, Y.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD2991
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-346 <KUR>
A:Cross-references: UNIPROT:Q8UA41; UNIPARC:UPI00000D2238; GB:AE008689; PIDN:AAL44346.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3534
A:Map position: linear chromosome
C:Superfamily: l-arabinose transport system permease araH

Query Match 5.1%; Score 6; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 242 LAALAA 247

RESULT 433
E98292
ribose ABC transporter (permease) rbsC AGR_L_2595 [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: E98292
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollan, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: E98292
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <KUR>
A:Cross-references: UNIPROT:Q8UA41; UNIPARC:UPI00000D2238; GB:AE007870; PIDN:AAK89863.1;
C:Genetics:
A:Gene: AGR_L_2595
A:Map position: linear chromosome
C:Superfamily: l-arabinose transport system permease araH

Query Match 5.1%; Score 6; DB 2; Length 346;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 242 LAALAA 247

RESULT 434
D83340
hypothetical protein PA2437 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: D83340
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
gen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: D83340
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-346 <STO>
A:Cross-references: UNIPROT:Q91145; UNIPARC:UPI00000C56D1; GB:AE004671; GB:AE004091; NID
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA2437

QY 21 LAALAA 26
|||
Db 207 LAALAA 212

RESULT 435

AI2974

tatrate dehydrogenase Atu3402 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C:Accession: AI2974
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo,
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AI2974
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-347 <KUR>
A:Cross-references: UNIPROT:Q8UAH2; UNIPARC:UPI000001647F9; GB:AE008689; PIDN:AAL44215.1
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu3402
A:Map position: linear chromosome
C:Superfamily: isocitrate dehydrogenase (NADP)

Query Match 5.1%; Score 6; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||
Db 342 VLAALA 347

RESULT 436

T02280

hypothetical protein T13D8.17 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 09-Jul-2004
C:Accession: T02280
R:Vysotskaia, V.S.; Schwartz, J.R.; Kwan, A.; Toriumi, M.; Yu, G.; Oji, O.; Liu, S.; Li,
itz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, J.
submitted to the EMBL Data Library, June 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC T13D8 sequence.
A:Reference number: Z14649
A:Accession: T02280
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-347 <VYS>
A:Cross-references: UNIPROT:O80751; UNIPARC:UPI00000A9321; EMBL:AC004473; NID:G3108025;
C:Genetics:
A:Gene: ATSP-T13D8.17
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein T13D8.25

Query Match 5.1%; Score 6; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GKGVLG 88
|||
Db 286 GKGVLG 291

RESULT 437

T14206

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 2 - Tanganicodus irsacae mitochondri
C:Species: mitochondrion Tanganicodus irsacae

C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
 C;Accession: T14206
 R;Kocher, T.D.; Conroy, J.A.; McKaye, K.R.; Stauffer, J.R.; Lockwood, S.F.
 Mol. Phylogenet. Evol. 4, 420-432, 1995
 A;Title: Evolution of NADH dehydrogenase subunit 2 in East African cichlid fish.
 A;Reference number: Z17790; MUID:96360498; PMID:8747298
 A;Accession: T14206
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-348 <KOC>
 A;Cross-references: UNIPROT:Q36057; UNIPARC:UPI000008EFD7; EMBL:U07265; NID:g463976; PID
 C;Experimental source: strain T13a; PSU
 C;Genetics:
 A;Genome: mitochondrion
 C;Superfamily: NADH dehydrogenase (ubiquinone) chain 2
 C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Query Match 5.1%; Score 6; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0;

Qy 21 LAALAA 26
 |||||
 Db 280 LAALAA 285

RESULT 438
 E87731
 protein F32B5.1 [imported] - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C;Accession: E87731
 R;Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998
 A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
 A;Reference number: A75000; MUID:99069613; PMID:9851916
 A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
 A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
 A;Accession: E87731
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-348 <STO>
 A;Cross-references: UNIPROT:O01854; UNIPARC:UPI00001756EC; GB:chr_I; PIDN:AAB54206.1; PI
 C;Genetics:
 A;Gene: F32B5.1
 A;Map position: 1
 C;Superfamily: creatine kinase; creatine kinase repeat homology

Query Match 5.1%; Score 6; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 GVLAAL 24
 |||||
 Db 44 GVLAAL 49

RESULT 439
 T29876
 hypothetical protein F32B5.1 - Caenorhabditis elegans (fragment)
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T29876
 R;Ledwith, J.; Graves, T.; Biewald, T.
 submitted to the EMBL Data Library, May 1997
 A;Description: The sequence of C. elegans cosmid F32B5.
 A;Reference number: Z20702
 A;Accession: T29876
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-348 <LED>
 A;Cross-references: UNIPROT:O01854; UNIPARC:UPI00001756EC; EMBL:AF003148; PIDN:AAB54206.
 A;Experimental source: strain Bristol N2; clone F32B5

C;Genetics:
 A;Gene: CESP:F32B5.1
 A;Map position: 1
 A;Introns: 61/3; 327/3
 C;Superfamily: creatine kinase; creatine kinase repeat homology

Query Match 5.1%; Score 6; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0;

Qy 19 GVLAAL 24
 |||||
 Db 44 GVLAAL 49

RESULT 440
 T03530
 cobW protein - Rhodobacter capsulatus
 C;Species: Rhodobacter capsulatus
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
 C;Accession: T03530
 R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
 Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
 A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003
 A;Reference number: Z14955; MUID:97404404; PMID:9256491
 A;Accession: T03530
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-348 <VLC>
 A;Cross-references: UNIPROT:O30787; UNIPARC:UPI00000B50E6; EMBL:AF010496; NID:g3128256;
 C;Genetics:
 A;Map position: 1
 C;Superfamily: cobW protein

Query Match 5.1%; Score 6; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 15 VLLGGV 20
 |||||
 Db 341 VLLGGV 346

RESULT 441
 A82429
 iron(III) ABC transporter, ATP-binding protein VCA0687 [imported] - Vibrio cholerae (str
 C;Species: Vibrio cholerae
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 31-Dec-2004
 C;Accession: A82429
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
 1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A;Reference number: A82035; MUID:20406833; PMID:10952301
 A;Accession: A82429
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-351 <HEI>
 A;Cross-references: UNIPROT:Q9KLQ5; UNIPARC:UPI00000C35F3; GB:AE004398; GB:AE003853; NID
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor
 C;Genetics:
 A;Gene: VCA0687
 A;Map position: 2

Query Match 5.1%; Score 6; DB 2; Length 351;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 6; Conservative 0;

Qy 5 ADLEVT 10
 |||||
 Db 261 ADLEVT 266

RESULT 442

D70788
probable trbB protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70788
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: D70788
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-352 <COL>
A: Cross-references: UNIPROT: O69627; UNIPARC: UPI00000D5EEP; GB: AL022121; GB: AL123456; NID
A: Experimental source: strain H37RV
C: Geneticks:
A: Gene: trbB
C: Superfamily: tumor-inducing plasmid pTrcS8 virB11 protein

Query Match 5.1%; Score 6; DB 2; Length 352;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 129 LAALAA 134

RESULT 443

D75460
MoxR-related protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: D75460
R: White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A: Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A: Reference number: A75250; MUID: 20036896; PMID: 10567266

A: Accession: D75460
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-354 <WHI>
A: Cross-references: UNIPROT: O9RVV4; UNIPARC: UPI00000D3D7D; GB: AE001944; GB: AE000513; NID
A: Experimental source: strain R1
C: Geneticks:
A: Gene: DR0918
A: Map position: 1
C: Superfamily: methanol dehydrogenase regulatory protein

Query Match 5.1%; Score 6; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 238 LAALAA 243

RESULT 444

A86843
prephenate dehydrogenase (EC 1.3.1.12) [imported] - Lactococcus lactis subsp. lactis (strain R1)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86843
R: Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001

A: Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A: Reference number: A86625; MUID: 21235186; PMID: 11337471
A: Accession: A86843
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-354 <STO>
A: Cross-references: UNIPROT: O9CET9; UNIPARC: UPI00001378A1; GB: AE005176; PID: g12724765;
A: Experimental source: strain IL1403
C: Geneticks:
A: Gene: tyrA
C: Superfamily: prephenate dehydrogenase, feedback inhibition-sensitive
C: Keywords: oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
|||||
Db 301 KVLGLL 306

RESULT 445

E83192
hypothetical protein PA3615 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: E83192
R: Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A: Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path
A: Reference number: A82950; MUID: 20437337; PMID: 10984043
A: Accession: E83192
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-356 <STO>
A: Cross-references: UNIPROT: O9HY11; UNIPARC: UPI00000C5A94; GB: AE004782; GB: AE004091; NI
A: Experimental source: strain PAO1
C: Geneticks:
A: Gene: PA3615

Query Match 5.1%; Score 6; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 27 LAALAA 32

RESULT 446

C69369
hypothetical protein AF0955 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: C69369
R: Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodso
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997

A: Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A: Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A: Reference number: A69250; MUID: 98049343; PMID: 9389475
A: Accession: C69369

A: Status: preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-359 <KLB>

A: Cross-references: UNIPROT: O29307; UNIPARC: UPI0000056EEA; GB: AE001038; GB: AE000782; NI
Query Match 5.1%; Score 6; DB 2; Length 359;

```

Best Local Similarity 100.0%; Pred. No. 5.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 83 KGKVLG 88
    |||||
Db 38 KGKVLG 43

RESULT 447
E72290
branched chain amino acid ABC transporter, permease protein - Thermotoga maritima (strain
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: E72290
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: E72290
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-359 <ARN>
A;Cross-references: UNIPROT:Q9XOM1; UNIPARC:UPI00000D3924; GB:AE001771; GB:AE000512; NID
A;Experimental source: strain W598
C;Genetics:
A;Gene: TM1137

Query Match 5.1%; Score 6; DB 2; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
    |||||
Db 118 GGVLA 123

RESULT 448
S74638
alanine dehydrogenase - Synecocystis sp. (strain PCC 6803)
N;Alternate names: hypothetical protein s111682
C;Species: Synecocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S74638
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S74638
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-360 <KAN>
A;Cross-references: UNIPROT:P72775; UNIPARC:UPI00000D347A; EMBL:D90900; GB:AB001339; NID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C;Superfamily: alanine dehydrogenase; alanine dehydrogenase homology
F11-277/Domain: alanine dehydrogenase homology <ALA>

Query Match 5.1%; Score 6; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 5.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
    |||||
Db 157 VLLGGV 162

RESULT 449
E75610
conserved hypothetical protein - Deinococcus radiodurans (strain R1)

```

```

C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: E75610
R;White, O.; Eissen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: E75610
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-361 <WHI>
A;Cross-references: UNIPROT:Q9RZ07; UNIPARC:UPI00000C1648; GB:AE001862; GB:AE001825; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0146
A;Map position: 2

Query Match 5.1%; Score 6; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 5.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLL 90
    |||||
Db 172 KVLGLL 177

RESULT 450
DB2644
sugar-phosphate dehydrogenase XF1723 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 05-Oct-2004
C;Accession: DB2644
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: AB2515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: DB2644
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-362 <SIM>
A;Cross-references: UNIPROT:Q9PCQ5; UNIPARC:UPI00000C27EC; GB:AE003996; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; F
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, F
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigi
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, I
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaas
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; 7
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1723
C;Superfamily: Aldehyde reductase

Query Match 5.1%; Score 6; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 5.3e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LIGGVL 21
    |||||
Db 256 LIGGVL 261

RESULT 451

```

T35079

hypothetical protein SCB126.09 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T36079
 R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, April 1999
 A:Reference number: Z21573
 A:Accession: T36079
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-362 <OLI>
 A:Cross-references: UNIPROT:Q9X852; UNIPARC:UPI000000DAFD6; EMBL:AL049630; PIDN:CAB40931.
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SCE126.09

Query Match 5.1%; Score 6; DB 2; Length 362;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
 |||||
 Db 93 GVLAAL 98

RESULT 452

AD2100
 alanine dehydrogenase [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AD2100
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD2100
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-363 <KUR>
 A:Cross-references: UNIPROT:Q8YU1; UNIPARC:UPI000000CE3B6; GB:BA000019; PIDN:BAE74054.1.
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr2355
 C:Superfamily: alanine dehydrogenase; alanine dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 363;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
 |||||
 Db 158 VLLGGV 163

RESULT 453

T34957
 Probable phospho-N-acetylmuramoyl-pentapeptide-transferase - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
 C:Accession: T34957
 R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, August 1999
 A:Reference number: Z21563
 A:Accession: T34957
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-363 <SAU>
 A:Cross-references: UNIPROT:P56833; UNIPARC:UPI000012F52B; EMBL:AL109663; PIDN:CAB51996.
 A:Experimental source: strain A3(2)
 C:Genetics:

A:Gene: SCOEDB:SC4A10.20C

C:Superfamily: phospho-N-acetylmuramoyl-pentapeptide-transferase

Query Match 5.1%; Score 6; DB 2; Length 363;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 271 LGGVLA 276

RESULT 454

G82734
 acetylornithine deacetylase XF1000 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 05-Oct-2004
 C:Accession: G82734
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: G82734
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-364 <SIM>
 A:Cross-references: UNIPROT:Q9PEM8; UNIPARC:UPI000000C25C0; GB:AE003938; GB:AE003849; N1
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Proh, J.D.; Junqueira, M.L.; Kemper, E.B.; Kitajima, J.P.; Krueger, J.E.; Kuramae, E.E.; Laig chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y., F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.V.; Palmieri, D. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawas, A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silvei M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF1000

C:Superfamily: Folate hydrolase G

Query Match 5.1%; Score 6; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 258 LAALAA 263

RESULT 455

S43574
 C05B5.3 protein (clone C05B5) - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
 C:Accession: S43574
 R:Mortimore, B.
 submitted to the EMBL Data Library, April 1994
 A:Reference number: S43570
 A:Accession: S43574
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-364 <MOR>
 A:Cross-references: UNIPARC:UPI000017B6AB; EMBL:Z32679
 C:Genetics:
 A:Introns: 1/1; 276/3; 294/3

Query Match 5.1%; Score 6; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EVTTST 13
|||||

Db 89 EVTTST 94
|||||

RESULT 456
C87292
conserved hypothetical protein CC0348 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: C87292
R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eissen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
N.; J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: C87292
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-364 <STO>
A/Cross-references: UNIPROT:Q9AB83; UNIPARC:UPI00000C6FF8; GB:AE005673; NID:gi3421499; E
C/Genetics:
A/Gene: CC0348

Query Match 5.1%; Score 6; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLALLA 25
|||||

Db 171 VLALLA 176
|||||

RESULT 457
A48135
transcription negative regulator AP-2B - human
C/Species: Homo sapiens (man)
C/Date: 30-Jun-1995 #sequence_revision 30-Jun-1995 #text_change 09-Jul-2004
C/Accession: A48135
R;Buettnar, R.; Kannan, P.; Imhof, A.; Bauer, R.; Yim, S.O.; Glockshuber, R.; Van Dyke,
Mol. Cell. Biol. 13, 4174-4185, 1993
A/Title: An alternatively spliced mRNA from the AP-2 gene encodes a negative regulator o
A/Reference number: A48135; MUID:93309451; PMID:8321221
A/Accession: A48135
A/Molecule type: mRNA
A/Residues: 1-365 <BUS>
A/Cross-references: UNIPROT:Q96GH0; UNIPARC:UPI0000178821; GB:M61156
C/Comment: This form does not dimerize and does not bind DNA. Its inhibition of the DNA
C/Genetics:
A/Gene: TrpA2
A/Map position: 6pter-p22.3
C/Superfamily: transcription factor AP-2
C/Keywords: alternative splicing; transcription regulation

Query Match 5.1%; Score 6; DB 2; Length 365;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||||

Db 248 LLGGVL 253
|||||

RESULT 458
F72062
hypothetical protein CP0185 [imported] - Chlamydophila pneumoniae (strains CWL029 and AP
C/Species: Chlamydophila pneumoniae; Chlamydia pneumoniae
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C/Accession: F72062; B81605
R;Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A/Reference number: A72000; MUID:99206606; PMID:10192388
A/Accession: F72062
A/Molecule type: DNA
A/Residues: 1-366 <ARN>
A/Cross-references: UNIPROT:Q9ZTV8; UNIPARC:UPI00000D4052; GB:AE001641; GB:AE001363; NII
A/Experimental source: strain CWL029
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A/Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39
A/Reference number: A81500; MUID:20150255; PMID:10684935
A/Accession: B81605
A/Molecule type: DNA
A/Residues: 1-366 <REA>
A/Cross-references: UNIPARC:UPI00000D4052; GB:AE002179; GB:AE002161; NID:g7189108; PIDN:
A/Experimental source: strain AR39, HL cells
C/Genetics:
A/Gene: CPn0565; CP0185
C/Superfamily: Chlamydia pneumoniae ct449 hypothetical protein

Query Match 5.1%; Score 6; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||

Db 40 LGGVLA 45
|||||

RESULT 459
A86561
CT449 hypothetical protein [imported] - Chlamydophila pneumoniae (strain J138)
C/Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C/Accession: A86561
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ie
Nucleic Acids Res. 28, 2311-2314, 2000
A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A/Reference number: A86491; MUID:20330349; PMID:10871362
A/Accession: A86561
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-366 <STO>
A/Cross-references: UNIPROT:Q9ZTV8; UNIPARC:UPI00000D4052; GB:BA000008; NID:g8978936; PJ
A/Experimental source: strain J138
C/Genetics:
A/Gene: CPJ0565
C/Superfamily: Chlamydia pneumoniae ct449 hypothetical protein

Query Match 5.1%; Score 6; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||

Db 40 LGGVLA 45
|||||

RESULT 460
S19172
cytochrome P450 2B4 - rat (fragments)
N/Alternate names: cytochrome P450 LM2
N/Contains: oxidoreductase (EC 1.-.-.-)
C/Species: Rattus norvegicus (Norway rat)
C/Date: 22-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C/Accession: S19172
R;Yuan, P.M.; Ryan, D.E.; Levin, W.; Shively, J.E.
Proc. Natl. Acad. Sci. U.S.A. 80, 1169-1173, 1983
A/Title: Identification and localization of amino acid substitutions between two phenobr

A:Reference number: S19172; MUID:83144040; PMID:6572377
A:Accession: S19172
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-158;159-200;201-310;311-367 <Y>A>
A:Cross-references: UNIPROT:Q7MOC4; UNIPARC:UPI0000174CD0; UNIPARC:UPI0000174CD1; UNIPARC:UPI0000174CD0;
A:Experimental source: strain Long-Evans
C:Genetics:
A:Gene: CYP2B4
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; heme; iron; metalloprotein; microsome; monooxygenase; oxidoreductase
P;312/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 AVIEPI 103
|||||
DB 102 AVIEPI 107

RESULT 461
AD0599
probable inner membrane protein STY0850 [imported] - Salmonella enterica subsp. enterica
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AD0599
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whittam, T.S.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Paratyphi C
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AD0599
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <PAR>
A:Cross-references: UNIPARC:UPI00005A0B3; GB:AL513382; PIDN:CAD05262.1; PID:gl6502029;
C:Genetics:
A:Gene: STY0850

Query Match 5.1%; Score 6; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 93 ATQOQA 98
|||||
DB 279 ATQOQA 284

RESULT 462
AI1079
conserved hypothetical protein lmo040 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AI1079
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, H.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; May, M.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, O.
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AI1079
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-369 <GLA>
A:Cross-references: UNIPROT:Q8YAS3; UNIPARC:UPI0000055682; GB:NC_003210; PIDN:CAC98255.1
A:Experimental source: strain EGD-e

Db 218 GVLAAL 223

RESULT 465

AC1272

alanine dehydrogenase homolog lml1579 [imported] - Listeria monocytogenes (strain EGD-e)

C/Species: Listeria monocytogenes

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AC1272

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A/Authors: Kretz, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AC1272

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-370 <GLA>

A/Cross-references: UNIPROT:Q8YGV2; UNIPARC:UPI0000054F86; GB:NC_003210; PIDN:CAC99657.1

A/Experimental source: strain EGD-e

C/Genetics:

A/Gene: lml1579

C/Superfamily: alanine dehydrogenase; alanine dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 370;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20

Db 157 VLLGGV 162

RESULT 466

AE1634

alanine dehydrogenase homolog lin1614 [imported] - Listeria innocua (strain Clip11262)

C/Species: Listeria innocua

C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AE1634

R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A/Authors: Kretz, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.

A/Reference number: AB1077; MUID:21537279; PMID:11679669

A/Accession: AE1634

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-370 <GLA>

A/Cross-references: UNIPROT:Q92BD6; UNIPARC:UPI00000CC5EA; GB:AL592022; PIDN:CAC96845.1

A/Experimental source: strain Clip11262

C/Genetics:

A/Gene: lin1614

C/Superfamily: alanine dehydrogenase; alanine dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 370;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20

Db 157 VLLGGV 162

RESULT 467

T10635

3-hydroxyisobutyrate dehydrogenase homolog T13K14.90 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cross)

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000

C/Accession: T10635

R/Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke submitted to the Protein Sequence Database, June 1999

A/Reference number: Z16991

A/Accession: T10635

A/Molecule type: DNA

A/Residues: 1-371 <BEV>

A/Cross-references: UNIPARC:UPI000016DACF; EMBL:AL080282; GSPDB:GN00062; ATSP:T13K14.90

A/Experimental source: cultivar Columbia; BAC clone T13K14

C/Genetics:

A/Gene: ATSP:T13K14.90

A/Map position: 4

A/Introns: 38/3; 68/2; 106/3; 157/1; 208/3; 240/3; 288/2; 316/3; 343/2

F:40-344/Domain: 3-hydroxyisobutyrate dehydrogenase; 3-hydroxyisobutyrate dehydrogenase h

Query Match 5.1%; Score 6; DB 2; Length 371;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23

Db 195 GGVLA 200

RESULT 468

B98308

3-isopropylmalate dehydrogenase [imported] - Agrobacterium tumefaciens (strain C58, Cere

C/Species: Agrobacterium tumefaciens

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 05-Oct-2004

C/Accession: B98308

R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A/Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: B98308

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-371 <KUR>

A/Cross-references: UNIPROT:Q8UAH2; UNIPARC:UPI00000D21BA; GB:AE007870; PIDN:AAK89988.1

C/Genetics:

A/Gene: AGR_L_2841

A/Map position: linear chromosome

C/Superfamily: isocitrate dehydrogenase (NADP)

Query Match 5.1%; Score 6; DB 2; Length 371;

Best Local Similarity 100.0%; Pred. No. 5.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

Db 366 VLAALA 371

RESULT 469

T31874

hypothetical protein F41B6.5 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C/Accession: T31874

R/Sammons, L.; Murray, J.

submitted to the EMBL Data Library, July 1997

A/Description: The sequence of C. elegans cosmid F41B6.

A/Reference number: Z21095

A/Accession: T31874

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-371 <SAM>

A/Cross-references: UNIPROT:O16457; UNIPARC:UPI0000164229; EMBL:AF016448; PIDN:AAB65955.1

A/Experimental source: strain Bristol N2; clone F41B6

C/Genetics:

A;Gene: CRSP:F41E6.5
A;Map position: 5
A;Introns: 47/2; 134/3; 230/2; 274/3; 327/3
C;Superfamily: alpha-hydroxy acid dehydrogenase, FMN-dependent; (S)-2-hydroxy-acid oxidase
P;4-309/Domain: (S)-2-hydroxy-acid oxidase homology <2HY>

Query Match 5.1%; Score 6; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
| | | | |
DB 335 VLGLLQ 340

RESULT 470
I51865
vasopressin V2 receptor - human
N;Alternate names: antidiuretic hormone receptor
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I51865; S24356
R;Seibold, A.; Brabet, P.; Rosenthal, W.; Birnbaumer, M.
Am. J. Hum. Genet. 51, 1078-1083, 1992
A;Title: Structure and chromosomal localization of the human antidiuretic hormone receptor
A;Reference number: I51865; MUID:93035372; PMID:1415251
A;Accession: I51865
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-371 <RES>
A;Cross-references: UNIPROT:P30519; UNIPARC:UPI000005043B; GB:I22206; NID:g347522; PIDN:
R;Birnbaumer, M.; Seibold, A.; Gilbert, S.; Ishido, M.; Barberis, C.; Antaramian, A.; Bix
Nature 357, 333-335, 1992
A;Title: Molecular cloning of the receptor for human antidiuretic hormone.
A;Reference number: S24356; MUID:92269957; PMID:1534149
A;Accession: S24356
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-371 <BIR>
A;Cross-references: UNIPARC:UPI000005043B; EMBL:Z11687; NID:g28417; PIDN:CAA77746.1; PID
C;Genetics:
A;Gene: GDB:AVPR2
A;Cross-references: GDB:I31475; OMIM:304800
A;Map position: Xq28-Xq28
A;Introns: 9/1; 304/1
C;Superfamily: oxytocin receptor
C;Keywords: G protein-coupled receptor; kidney; transmembrane protein

Query Match 5.1%; Score 6; DB 2; Length 371;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
| | | | |
DB 58 VLAALA 63

RESULT 471
AD0262
probable membrane protein YPO2151 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD0262
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0262
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <KUR>

A;Cross-references: UNIPROT:Q8ZEL3; UNIPARC:UPI00000DCE84; GB:AL590842; PIDN:CAC90960.1
C;Genetics:
A;Gene: YPO2151

Query Match 5.1%; Score 6; DB 2; Length 373;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|
|
|
|
Db 12 GVLAAL 17

RESULT 472
S48639
fructose-bisphosphate aldolase (EC 4.1.2.13) precursor - Chlamydomonas reinhardtii
N;Alternate names: fructose-1,6-bisphosphate aldolase
C;Species: Chlamydomonas reinhardtii
C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jul-2004
C;Accession: S48639; S58485; S58486; S34367
R;Schnarrenberger, C.; Pelzer-Reith, B.; Yatsuki, H.; Freund, S.; Jacobshagen, S.; Hori, A.; Title: Expression and sequence of the only detectable aldolase in Chlamydomonas reinhardtii: intron/exon organization
A;Reference number: S58485; MUID:96004769; PMID:7565612
A;Accession: S58485
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 28-374 <PEL>
A;Cross-references: UNIPARC:UPI0000175FF0; EMBL:X85495
R;Pelzer-Reith, B.
submitted to the EMBL Data Library, March 1995
A;Reference number: S58486
A;Accession: S58486
A;Molecule type: DNA
A;Residues: 1-367,370-374 <PEW>
A;Cross-references: UNIPARC:UPI0000175FF1; EMBL:X85495
C;Genetics:
A;Gene: ALDCHL
A;Genome: nuclear
A;Introns: 209/3; 235/3; 325/3
C;Superfamily: fructose-bisphosphate aldolase
C;Keywords: aldehyde-lyase; carbon-carbon lyase; chloroplast
F;1-27/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;28-374/Product: fructose-bisphosphate aldolase #status predicted <MAT>
F;162,243,374/Active site: Lys, Lys, Tyr #status predicted

Query Match 5.1%; Score 6; DB 2; Length 374;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|
|
|
|
Db 16 LAALAA 21

RESULT 473
F87596
hypothetical protein CC2806 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: F87596
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

J. Mol. Biol. 156, 683-717, 1982
A;Title: Complete sequence of bovine mitochondrial DNA. Conserved features of the mammal
A;Reference number: A00152; MUID:83010260; PMID:7120390
A;Accession: A00152
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P00157; UNIPARC:UPI0000001E76; GB:J01394; NID:g336430; PIDN:
J. Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S27087
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPARC:UPI0000001E76
A;Note: this sequence and translation are not annotated in GenBank release 111.0, but do
J. He, D.Y.; Yu, L.; Yu, C.A.
J. Biol. Chem. 269, 2292-2298, 1994
A;Title: Ubiquinone binding domains in bovine heart mitochondrial cytochrome b.
A;Reference number: A49734; MUID:94124591; PMID:8294488
A;Accession: A49734
A;Molecule type: protein
A;Residues: 142-146;327-332 <HEA>
A;Cross-references: UNIPARC:UPI0000171CE4; UNIPARC:UPI0000171CE5
R. Xie, D.; Yu, C.A.; Kim, H.; Xia, J.Z.; Kachurin, A.M.; Zhang, L.; Yu, L.; Deisenhofer, R.
Science 277, 60-66, 1997
A;Title: Crystal structure of the cytochrome b_c1 complex from bovine heart mitochondria.
A;Reference number: A58900; MUID:97349328; PMID:9204897
A;Contents: annotation; X-ray crystallography, 2.9 angstroms
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Complex: the transmembrane complex includes cytochrome b, cytochrome c1 (see PIR:CCBO1
24864, PIR:ZPB0C1, and PIR:ZPB0C2)
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b6 homology <CBH>
F;11-209/Domain: transmembrane #status predicted <TM1>
F;11-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;228-304/Domain: transmembrane #status predicted <TM5>
F;323-343/Domain: transmembrane #status predicted <TM6>
F;353-369/Domain: transmembrane #status predicted <TM7>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b6 homology <CBH>
F;11-209/Domain: transmembrane #status predicted <TM1>
F;11-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;228-304/Domain: transmembrane #status predicted <TM5>
F;323-343/Domain: transmembrane #status predicted <TM6>
F;353-369/Domain: transmembrane #status predicted <TM7>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status experime
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status experime

A;Accession: S17405
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24992; UNIPARC:UPI00000001F0; EMBL:X56286; NID:g12624; PIDN:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b6 homology <CBH>
F;11-209/Domain: transmembrane #status predicted <TM1>
F;11-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;228-304/Domain: transmembrane #status predicted <TM5>
F;323-343/Domain: transmembrane #status predicted <TM6>
F;353-369/Domain: transmembrane #status predicted <TM7>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted
Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LGGVLA 22
|||||
DB 288 LGGVLA 293
RESULT 483
S17406
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Arabian camel mitochondrion
C;Species: mitochondrion Camelus dromedarius (Arabian camel)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
A;Accession: S17406
R. Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17406
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24952; UNIPARC:UPI000012895E; EMBL:X56281; NID:g12854; PIDN:
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b6 homology <CBH>
F;11-209/Domain: transmembrane #status predicted <TM1>
F;11-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;228-304/Domain: transmembrane #status predicted <TM5>
F;323-343/Domain: transmembrane #status predicted <TM6>
F;353-369/Domain: transmembrane #status predicted <TM7>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

P;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 484
S17407
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - goat mitochondrion
C:Species: mitochondrion Capra aegagrus hircus (domestic goat)
C>Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C:Accession: S17407
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A>Title: Evolution of the cytochrome b gene of mammals.
A:Reference number: S17405; MUID:91178817; PMID:1901092
A:Accession: S17407
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <IRW>
A:Cross-references: UNIPROT:P24953; UNIPARC:UPI0000128965; EMBL:X56289; NID:g12871; PIDN
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Function:
A>Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released
A:Pathway: oxidative phosphorylation; respiratory chain
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocytin
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocytin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 485
S17408
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - black rhinoceros mitochondrion
C:Species: mitochondrion Diceror bicornis (black rhinoceros)
C>Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C:Accession: S17408
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A>Title: Evolution of the cytochrome b gene of mammals.
A:Reference number: S17405; MUID:91178817; PMID:1901092
A:Accession: S17408
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <IRW>
A:Cross-references: UNIPROT:P24954; UNIPARC:UPI00001289A5; EMBL:X56283; NID:g12903; PIDN

C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Function:
A>Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released
A:Pathway: oxidative phosphorylation; respiratory chain
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocytin
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocytin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 486
S17409
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - fallow deer mitochondrion
C:Species: mitochondrion Dama dama (fallow deer)
C>Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C:Accession: S17409
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A>Title: Evolution of the cytochrome b gene of mammals.
A:Reference number: S17405; MUID:91178817; PMID:1901092
A:Accession: S17409
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <IRW>
A:Cross-references: UNIPROT:P24955; UNIPARC:UPI000016D507; EMBL:X56290; NID:g12907; PIDN
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Function:
A>Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released
A:Pathway: oxidative phosphorylation; respiratory chain
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocytin
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocytin-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: transmembrane #status predicted <TM5>
F;229-245/Domain: transmembrane #status predicted <TM6>
F;288-304/Domain: transmembrane #status predicted <TM7>
F;323-343/Domain: transmembrane #status predicted <TM8>
F;353-369/Domain: transmembrane #status predicted <TM9>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 490
S17414
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - mule deer mitochondrion
C;Species: mitochondrion Odocoileus hemionus (mule deer)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17414
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17414
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24960; UNIPARC:UPI0000128A6B; EMBL:X56291; NID:g13198; PIDN
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17414
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24960; UNIPARC:UPI0000128A6B; EMBL:X56291; NID:g13198; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastocyanin
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: transmembrane #status predicted <TM5>
F;229-245/Domain: transmembrane #status predicted <TM6>
F;288-304/Domain: transmembrane #status predicted <TM7>
F;323-343/Domain: transmembrane #status predicted <TM8>
F;353-369/Domain: transmembrane #status predicted <TM9>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 491
S17417
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b, isolate 1B - pantropical sp

C;Species: mitochondrion Stenella longirostris (pantropical spinner dolphin)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
C;Accession: S17417; S17416
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17417
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24962; UNIPARC:UPI00000001E0; EMBL:X56293; NID:g13628; PIDN
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17417
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPARC:UPI0000128B12; EMBL:X56292; NID:g13626; PIDN:CAA39739.1; PI
A;Experimental source: isolate 1B
A;Accession: S17416
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-59, 'T', '61-97, 'M', '99-265, 'P', '267-299, 'I', '301-326, 'V', '328-379 <IR2>
A;Cross-references: UNIPARC:UPI0000128B12; EMBL:X56292; NID:g13626; PIDN:CAA39739.1; PI
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastocyanin
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: transmembrane #status predicted <TM5>
F;229-245/Domain: transmembrane #status predicted <TM6>
F;288-304/Domain: transmembrane #status predicted <TM7>
F;323-343/Domain: transmembrane #status predicted <TM8>
F;353-369/Domain: transmembrane #status predicted <TM9>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 492
S17418
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - pig mitochondrion
C;Species: mitochondrion Sus scrofa domestica (domestic pig)
C;Date: 29-Jan-1993 #sequence_revision 23-Jul-1999 #text_change 09-Jul-2004
C;Accession: T10984; T11882; S17418; S58080; S58021; S58059; S58058; S58060; S58079
R;Lin, C.S.; Lin, C.Y.; Sun, Y.L.; Chang, L.C.; Cheng, I.C.; Yang, P.C.; Mao, S.J.T.; H
submitted to the EMBL Data Library, November 1997
A;Description: Complete nucleotide sequence of the porcine mitochondrial genome.
A;Reference number: Z17237
A;Accession: T10984
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-379 <LIN>
A;Cross-references: UNIPROT:P24964; UNIPARC:UPI00000011F0; EMBL:AF034253; NID:g4958951;
R;Ursing, B.M.
submitted to the EMBL Data Library, February 1999
A;Description: The complete mitochondrial DNA sequence of the pig (Sus scrofa).
A;Reference number: Z17370
A;Accession: T11882
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-313,'G','315-379 <URS>
A;Cross-references: UNIPARC:UPI0000000806; EMBL:AJ002189; PIDN:CAA05239.1
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17418
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-197,'M',199-313,'G',315-379 <IRW>
A;Cross-references: UNIPARC:UPI000016D652; EMBL:X56295; NID:g13678; PIDN:CAA39742.1; PID
R;Randi, E.; Lucchini, V.; Dong, C.
submitted to the EMBL Data Library, July 1995
A;Description: Evolutionary genetics of the suiformes.
A;Reference number: S58021
A;Accession: S58080
A;Molecule type: DNA
A;Residues: 1-379 <RAW>
A;Cross-references: UNIPARC:UPI00000011F0; EMBL:Z50087; NID:g902692; PIDN:CAA90418.1; PI
A;Experimental source: subspecies leucomystax
A;Accession: S58021
A;Molecule type: DNA
A;Residues: 1-379 <RAW>
A;Cross-references: UNIPARC:UPI00000011F0; EMBL:Z50087; NID:g902692; PIDN:CAA90418.1; PI
A;Experimental source: subspecies Asian domestic pig
A;Accession: S58059
A;Molecule type: DNA
A;Residues: 1-313,'G','315-379 <RA2>
A;Cross-references: UNIPARC:UPI0000000806; EMBL:Z50088; NID:g902696; PIDN:CAA90419.1; PI
A;Experimental source: subspecies majori
A;Accession: S58058
A;Molecule type: DNA
A;Residues: 1-313,'G','315-379 <RA3>
A;Cross-references: UNIPARC:UPI0000000806; EMBL:Z50085; NID:g902698; PIDN:CAA90416.1; PI
A;Experimental source: subspecies lybicus, isolate Bulgarian (1)
A;Accession: S58060
A;Molecule type: DNA
A;Residues: 1-313,'G','315-367','V',369-379 <RA4>
A;Cross-references: UNIPARC:UPI00000001F1; EMBL:Z50089; NID:g902694; PIDN:CAA90420.1; PI
A;Experimental source: subspecies meridionalis (Sardinian wild boar)
A;Accession: S58079
A;Molecule type: DNA
A;Residues: 1-294,'M',296-313,'G',315-379 <RA5>
A;Cross-references: UNIPARC:UPI000016D8A7; EMBL:Z50086; NID:g902690; PIDN:CAA90417.1; PI
A;Experimental source: subspecies lybicus, isolate Bulgarian (2)
C;Genetics:
A;Gene: cytb
A;Genome: mitochondrion
A;Genetic code: SGC1
A;Note: cytb
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;117-133/Domain: transmembrane #status predicted <TM1>
F;178-200/Domain: transmembrane #status predicted <TM2>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 493

S17419
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balabac chevrotain mitoch
C;Species: mitochondrion Tragulus napu (Balabac chevrotain)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
A;Accession: S17419
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17419
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24965; UNIPARC:UPI0000128B29; EMBL:X56288; NID:g13836; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;136-52/Domain: transmembrane #status predicted <TM1>
F;181-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 494

S17420
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - collared peccary mitochond
C;Species: mitochondrion Tayassu tajacu (collared peccary)
C;Date: 29-Jan-1993 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
A;Accession: S17420
R;Irwin, D.M.; Kocher, T.D.; Wilson, A.C.
J. Mol. Evol. 32, 128-144, 1991
A;Title: Evolution of the cytochrome b gene of mammals.
A;Reference number: S17405; MUID:91178817; PMID:1901092
A;Accession: S17420
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <IRW>
A;Cross-references: UNIPROT:P24966; UNIPARC:UPI0000128B1E; EMBL:X56296; NID:g13874; PIDN
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released.
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 495
 S26163
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - harbor seal mitochondrion
 C:Species: mitochondrion Phoca vitulina (harbor seal)
 C:Date: 03-Feb-1994 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C:Accession: S26163; S58453
 R:Arnasson, U.; Johnsson, E.
 J. Mol. Evol. 34, 493-505, 1992
 A:Title: The complete mitochondrial DNA sequence of the harbor seal, Phoca vitulina.
 A:Reference number: S26153; MUID:92277666; PMID:1593642
 A:Accession: S26163
 A:Molecule type: DNA
 A:Residues: 1-379 <ARN>
 A:Cross-references: UNIPROT:Q00530; UNIPARC:UPI0000163909; EMBL:X63726; NID:g13431; PIDN
 R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
 J. Mol. Evol. 40, 78-85, 1995
 A:Title: A molecular view of pinned relationships with particular emphasis on the true
 A:Reference number: S58447; MUID:95230701; PMID:7714914
 A:Accession: S58453
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-14, 'D', 16-189, 'S', 191-192, 'A', 194-213, 'N', 215-258, 'A', 260-299, 'I', 301-303,
 A:Cross-references: UNIPARC:UPI0000128A90; EMBL:X82306; NID:g693981; PIDN:CAA57749.1; PI
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGCI
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:36-52/Domain: cytochrome b6 homology <CB6>
 F:81-99/Domain: transmembrane #status predicted <TM1>
 F:117-133/Domain: transmembrane #status predicted <TM2>
 F:178-200/Domain: transmembrane #status predicted <TM3>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 496
 S33572
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - southern African porcupin.
 C:Species: mitochondrion Hystrix africaeaustralis (southern African porcupine)
 C:Date: 03-Feb-1994 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C:Accession: S33572
 R:Ma, D.P.; Zharkikh, A.; Graur, D.; VandeBerg, J.L.; Li, W.H.
 J. Mol. Evol. 36, 327-334, 1993
 A:Title: Structure and evolution of opossum, guinea pig, and porcupine cytochrome b ge
 A:Reference number: S33572; MUID:93301932; PMID:8315653
 A:Accession: S33572
 A:Molecule type: DNA
 A:Residues: 1-379 <MAD>
 A:Cross-references: UNIPROT:Q04910; UNIPARC:UPI000016D69D; EMBL:X70674; NID:g14012; PID
 A:Note: residue 1 and the corresponding nucleotide sequence are not shown
 C:Genetics:
 A:Gene: cob
 A:Genome: mitochondrion
 A:Genetic code: SGCI
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase compl
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rel
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:36-52/Domain: cytochrome b6 homology <CB6>
 F:81-99/Domain: transmembrane #status predicted <TM1>
 F:117-133/Domain: transmembrane #status predicted <TM2>
 F:178-200/Domain: transmembrane #status predicted <TM3>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 497
 S41832
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - blue whale mitochondrion
 C:Species: mitochondrion Balaenoptera musculus (blue whale)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C:Accession: S41832
 R:Arnasson, U.; Gullberg, A.
 J. Mol. Evol. 37, 312-322, 1993
 A:Title: Comparison between the complete mtDNA sequences of the blue and the fin whale.
 A:Reference number: S41820; MUID:94141932; PMID:8308901
 A:Accession: S41832
 A:Molecule type: DNA
 A:Residues: 1-379 <ARN>
 A:Cross-references: UNIPROT:P41285; UNIPARC:UPI0000128949; EMBL:X72204; NID:g414126; PI
 C:Genetics:


```
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 498
S41833
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Weddell seal mitochondrion
C:Species: mitochondrion Lepconychotes weddelli (Weddell seal)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S41833
R:Arnason, U.; Gullberg, A.; Johnsson, E.; Ledje, C.
J. Mol. Evol. 37, 323-330, 1993
A:Title: The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, Hal
A:Reference number: S41833; MUID:94141933; PMID:8308902
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P38594; UNIPARC:UPI0000128A1B; EMBL:X72005; NID:9414771; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 499
S41834
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Monachus schauinslandi mit
C:Species: mitochondrion Monachus schauinslandi
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S41834
R:Arnason, U.; Gullberg, A.; Johnsson, E.; Ledje, C.
J. Mol. Evol. 37, 323-330, 1993
A:Title: The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, Hal
A:Reference number: S41833; MUID:94141933; PMID:8308902
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P38591; UNIPARC:UPI000009033C; EMBL:X72209; NID:9414773; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 500
S41847
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - gray seal mitochondrion
C:Species: mitochondrion Halichoerus grypus (gray seal)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S41847
R:Arnason, U.; Gullberg, A.; Johnsson, E.; Ledje, C.
J. Mol. Evol. 37, 323-330, 1993
A:Title: The nucleotide sequence of the mitochondrial DNA molecule of the grey seal, Hal
A:Reference number: S41833; MUID:94141933; PMID:8308902
A:Status: nucleic acid sequence not shown; translation not shown
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P38593; UNIPARC:UPI00001289EC; EMBL:X72004; NID:9414757; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1993
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
```


F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 501
S43261
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - minke whale mitochondrion
C:Species: mitochondrion Balaenoptera acutorostrata (minke whale, lesser rorqual)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S43261
R:Arnasson, U.; Gullberg, A.
Nature 367, 726-728, 1994
A:Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
A:Reference number: S43261; MUID:94150700; PMID:8107866
A:Accession: S43261
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P41280; UNIPARC:UPI0000128944; EMBL:X75753; NID:9457763; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; in

F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 502
S43262
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaenoptera bonaerensis
C:Species: mitochondrion Balaenoptera bonaerensis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S43262
R:Arnasson, U.; Gullberg, A.
Nature 367, 726-728, 1994
A:Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
A:Reference number: S43261; MUID:94150700; PMID:8107866
A:Accession: S43262

A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P41281; UNIPARC:UPI0000128945; EMBL:X75581; NID:9457762; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; in

F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 503
S43263
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - sei whale mitochondrion
C:Species: mitochondrion Balaenoptera borealis (sei whale)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S43263
R:Arnasson, U.; Gullberg, A.
Nature 367, 726-728, 1994
A:Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
A:Reference number: S43261; MUID:94150700; PMID:8107866
A:Accession: S43263
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P41282; UNIPARC:UPI0000128946; EMBL:X75582; NID:9457763; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; in

F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

QY 17 LGGVLA 22
 Db 288 LGGVLA 293
 |||||
 |||||

RESULT 504

S43264
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Bryde's whale mitochondrion
 C:Species: mitochondrion Balaenoptera edeni (Bryde's whale)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 A:Accession: S43264
 R:Arnason, U.; Gullberg, A.
 Nature 367, 726-728, 1994
 A>Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
 A:Reference number: S43261; MUID:94150700; PMID:8107866
 A:Accession: S43264
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <ARN>
 A:Cross-references: UNIPROT:P41283; UNIPARC:UPI0000128947; EMBL:X75583; NID:g457766; PID
 C:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
 in

F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: transmembrane #status predicted <TM1>
 F:36-52/Domain: transmembrane #status predicted <TM2>
 F:81-99/Domain: transmembrane #status predicted <TM3>
 F:117-133/Domain: transmembrane #status predicted <TM4>
 F:178-200/Domain: transmembrane #status predicted <TM5>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM6>
 F:288-304/Domain: transmembrane #status predicted <TM7>
 F:323-343/Domain: transmembrane #status predicted <TM8>
 F:353-369/Domain: transmembrane #status predicted <TM9>
 F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 Db 288 LGGVLA 293
 |||||
 |||||

RESULT 505

S43265
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaena glacialis mitochondrion
 C:Species: mitochondrion Balaena glacialis
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 A:Accession: S43265
 R:Arnason, U.; Gullberg, A.
 Nature 367, 726-728, 1994
 A>Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
 A:Reference number: S43261; MUID:94150700; PMID:8107866
 A:Accession: S43265
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <ARN>
 A:Cross-references: UNIPROT:P41284; UNIPARC:UPI0000171CB6; EMBL:X75587
 C:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
 in

F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 Db 288 LGGVLA 293
 |||||
 |||||

RESULT 506

S43266
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Balaena mysticetus mitochondrion
 C:Species: mitochondrion Balaena mysticetus
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 A:Accession: S43266
 R:Arnason, U.; Gullberg, A.
 Nature 367, 726-728, 1994
 A>Title: Relationship of baleen whales established by cytochrome b gene sequence comparison
 A:Reference number: S43261; MUID:94150700; PMID:8107866
 A:Accession: S43266
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <ARN>
 A:Cross-references: UNIPROT:P41286; UNIPARC:UPI000012894A; EMBL:X75588; NID:g457770; PID
 C:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
 in

F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 Db 288 LGGVLA 293
 |||||
 |||||

RESULT 507

S43267
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Caperea marginata mitochondrion
 C:Species: mitochondrion Caperea marginata
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted;
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted;

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||
Db 288 LGGVLA 293

RESULT 509
S43269
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - humpback whale mitochondr-
C/Species: mitochondrion Megaptera novaeangliae (humpback whale)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S43269
R/Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A/Title: Relationship of baleen whales established by cytochrome b gene sequence compari-
A/Reference number: S43261; MUID:94150700; PMID:8107866
A/Accession: S43269
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-379 <ARN>
A/Cross-references: UNIPROT:P41289; UNIPARC:UPI0000128A3A; EMBL:X75584; NID:9457794; PTI:
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGCI
C/Suprafamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol-
C/Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; in
in
F/11-339/Domain: cytochrome b homology <CBH>
F/11-209/Domain: cytochrome b6 homology <CB6>
F/36-52/Domain: transmembrane #status predicted <TM1>
F/81-99/Domain: transmembrane #status predicted <TM2>
F/117-133/Domain: transmembrane #status predicted <TM3>
F/178-200/Domain: transmembrane #status predicted <TM4>
F/221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F/229-245/Domain: transmembrane #status predicted <TM5>
F/288-304/Domain: transmembrane #status predicted <TM6>
F/323-343/Domain: transmembrane #status predicted <TM7>
F/353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted;
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted;

Query Match 5.1%; Score 6; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||
Db 288 LGGVLA 293

RESULT 510
S43270
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Physeter macrocephalus mit-
C/Species: mitochondrion Physeter macrocephalus
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S43270
R/Arnason, U.; Gullberg, A.
Nature 367, 726-728, 1994
A/Title: Relationship of baleen whales established by cytochrome b gene sequence compari-
A/Reference number: S43261; MUID:94150700; PMID:8107866
A/Accession: S43270
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-379 <ARN>
A/Cross-references: UNIPROT:P41290; UNIPARC:UPI0000000808; EMBL:X75589; NID:9457797; PTI:
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

R:Janke, A.; Gemmell, N.J.; Feldmaier-Fuchs, G.; von Haeseler, A.; Paabo, S.
 J. Mol. Evol. 42, 153-159, 1996
 A:Title: The mitochondrial genome of a monotreme--the platypus (*Ornithorhynchus anatinus*)
 A:Reference number: A58888; MUID:97077300; PMID:8919867
 A:Accession: E58889
 A:Status: nucleic acid sequence not shown; translation not shown; not compared with cont
 A:Molecule type: DNA
 A:Residues: 1-379 <JAN>
 A:Cross-references: UNIPROT:Q36461; UNIPARC:UPI0000128A72; GB:X83427; NID:g1469249; PIDN
 A:Note: submitted to GenBank/EMBL/DBJ December, 1994
 C:Genetics:
 A:Map position: FOR14186-15325
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:142-146/Region: ubiquinone binding #status predicted <TM4>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 515
 I48133
 ubi quinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl
 C:Species: mitochondrion Dactylomys boliviensis (American spiny rat)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
 C:Accession: I48133
 R:da Silva, M.N.F.; Patton, J.L.
 Mol. Phylogenet. Evol. 2, 243-255, 1993
 A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodents
 A:Reference number: A49605; MUID:94184505; PMID:8136924
 A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodents
 A:Reference number: A49605; MUID:94184505; PMID:8136924
 A:Accession: I48133
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-379 <RES>
 A:Cross-references: UNIPROT:Q34301; UNIPARC:UPI0000095C92; GB:L23339; NID:g996082; PIDN:
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 515
 I48133
 ubi quinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl
 C:Species: mitochondrion Dactylomys boliviensis (American spiny rat)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
 C:Accession: I48133
 R:da Silva, M.N.F.; Patton, J.L.
 Mol. Phylogenet. Evol. 2, 243-255, 1993
 A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodents
 A:Reference number: A49605; MUID:94184505; PMID:8136924
 A:Accession: I48133
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-379 <RES>
 A:Cross-references: UNIPROT:Q34301; UNIPARC:UPI0000095C92; GB:L23339; NID:g996082; PIDN:
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 516
 I48132
 ubi quinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Dactyl
 C:Species: mitochondrion Dactylomys dactylinus (American spiny rat)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
 C:Accession: I48132
 R:da Silva, M.N.F.; Patton, J.L.
 Mol. Phylogenet. Evol. 2, 243-255, 1993
 A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodent
 A:Reference number: A49605; MUID:94184505; PMID:8136924
 A:Accession: I48132
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-379 <RES>
 A:Cross-references: UNIPROT:Q34306; UNIPARC:UPI0000092E09; GB:L23335; NID:g995850; PIDN:
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 288 LGGVLA 293

RESULT 517
 I48135
 ubi quinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Bchim
 C:Species: mitochondrion Echimy's didelphoides (American spiny rat)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
 C:Accession: I48135; I48136; I48137; I48138
 R:da Silva, M.N.F.; Patton, J.L.
 Mol. Phylogenet. Evol. 2, 243-255, 1993
 A:Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyd rodent
 A:Reference number: A49605; MUID:94184505; PMID:8136924
 A:Accession: I48135
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EM
 A:Molecule type: DNA
 A:Residues: 1-379 <SIL1>
 A:Cross-references: UNIPROT:Q34424; UNIPARC:UPI00000900D3; GB:L23363; NID:g996078; PIDN:
 A:Experimental source: individual LHE 600
 A:Note: the source is designated as Makalata didelphoides
 A:Accession: I48136
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EM
 A:Molecule type: DNA
 A:Residues: 1-59, 'Y', 61-78, 'I', 80-107, 'S', 109-163, 'V', 165-179, 'A', 181, 'X', 183-204, 'X', 2
 A:Cross-references: UNIPARC:UPI0000093B20; GB:L23362; NID:g995570; PIDN:AAC37692.1; PID:
 A:Experimental source: individual LHE 595
 A:Accession: I48137
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EM
 A:Molecule type: DNA
 A:Residues: 1-22, 'T', 24-28, 'T', 30-59, 'Y', 61-78, 'L', 80-81, 'T', 83-88, 'L', 90-114, 'I', 116-12
 A:Cross-references: UNIPARC:UPI0000090C47; GB:L23359; NID:g995842; PIDN:AAC37690.1; PID:
 A:Experimental source: individual INPA 2474
 A:Accession: I48138
 A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EM
 A:Molecule type: DNA
 A:Residues: 1-17, 'L', 19-22, 'T', 24-28, 'T', 30-59, 'Y', 61-78, 'L', 80-81, 'T', 83-88, 'L', 90-114, 'I', 116-12
 A:Cross-references: UNIPARC:UPI000008CC1C; GB:L23356; NID:g995844; PIDN:AAC37688.1; PID:
 A:Experimental source: individual LHE 595

C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F;11-339/Domain: cytochrome b homology <CYB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: cytochrome b6 homology <CB6>
F;83,182/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 522
T11505
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - fruit bat (Artibeus jamaicensis)
C;Species: mitochondrion Artibeus jamaicensis
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11505
R;Pumo, D.E.; Finamore, P.S.; Franek, W.R.; Phillips, C.J.; Tarzami, S.; Balzarano, D.
J. Mol. Evol. 47, 709-717, 1998
A;Title: Complete mitochondrial genome of a neotropical fruit bat, Artibeus jamaicensis,
A;Reference number: Z17251; MUID:99065784; PMID:9847413
A;Accession: T1152
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-379 <PUM>
A;Cross-references: UNIPROT:Q99606; UNIPARC:UPI000008B486; EMBL:AF061340; NID:94164474;
C;Genetics:
A;Genome: mitochondrion
A;Note: Cytb
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F;11-339/Domain: cytochrome b homology <CYB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: cytochrome b6 homology <CB6>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 523
T11505
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - dog mitochondrion
C;Species: mitochondrion Canis lupus familiaris (dog)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: T11505
R;Kim, K.S.; Lee, S.E.; Jeong, H.W.; Ha, J.H.
Mol. Phylogenet. Evol. 10, 210-220, 1998
A;Title: The complete nucleotide sequence of the domestic dog (Canis familiaris) mitochondrion
A;Reference number: Z17276; MUID:99097286; PMID:9878232
A;Accession: T11505
A;Status: nucleic acid sequence not shown; translated from GB/EMBL
A;Molecule type: DNA
A;Residues: 1-379 <KIM>
A;Cross-references: UNIPROT:Q34101; UNIPARC:UPI000000A03D; EMBL:U96639; NID:94154170; PT
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1

A;Note: cytb
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F;11-339/Domain: cytochrome b homology <CYB>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: cytochrome b6 homology <CB6>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 524
S58455
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American black bear mitochondrion
C;Species: mitochondrion Ursus americanus (American black bear)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S58455
R;Arnason, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A;Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A;Reference number: S58447; MUID:95230701; PMID:7714914
A;Accession: S58455
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-379 <ARN>
A;Cross-references: UNIPROT:Q9G104; UNIPROT:Q8SJ17; UNIPROT:Q9G4V2; UNIPROT:Q9G4V3; UNIPROT:Q9G4V4
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion
F;11-339/Domain: cytochrome b homology <CB6>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: cytochrome b6 homology <CB6>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 525
S58454
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - brown bear mitochondrion
C;Species: mitochondrion Ursus arctos (brown bear)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S58454
R;Arnason, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A;Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A;Reference number: S58447; MUID:95230701; PMID:7714914

Qy 17 LGGVLA 22
|
|
|
|
Db 288 LGGVLA 293

RESULT 530
S58466
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - walrus mitochondrion
C:Species: mitochondrion Odobenus rosmarus (walrus)
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S58466
E:Arnaason, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58466
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:CROSS-references: UNIPROT:Q37769; UNIPARC:UPI0000093321; EMBL:X82299; NID:g693967; PIR
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion,
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83, 182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97, 196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted.

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|
|
|
|
Db 288 LGGVLA 293

RESULT 531
S58465
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - southern elephant seal mit
C:Species: mitochondrion Mirounga leonina (southern elephant seal)
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S58465
E:Arnaason, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58465
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:CROSS-references: UNIPROT:Q35019; UNIPARC:UPI0000128A43; EMBL:X82298; NID:g693965; PIR
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion,
F:11-339/Domain: cytochrome b homology <CBH>

F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:228-245/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 532
ES5849
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - ringed seal mitochondrion
C:Species: mitochondrion Phoca hispida, Pusa hispida (ringed seal)
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S58449
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58449
A:Status: nucleic acid sequence not shown; translation not shown
A:Keywords: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q35468; UNIPARC:UPI0000128A8D; EMBL:X82304; NID:G693973; PID:G693973
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-209/Domain: cytochrome b homology <CB6>
F:36-52/Domain: cytochrome b6 homology <CB6>
F:81-99/Domain: transmembrane #status predicted <TM1>
F:117-133/Domain: transmembrane #status predicted <TM2>
F:142-146/Region: ubiquinone binding #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:228-245/Domain: transmembrane #status predicted <TM5>
F:323-343/Domain: transmembrane #status predicted <TM6>
F:353-369/Domain: transmembrane #status predicted <TM7>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 533
ES5851
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - finback whale mitochondrion
C:Species: mitochondrion Balaenoptera physalus (finback whale, common rorqual)
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 09-Jul-2004

C:Accession: ES5851
R:Arnasson, U.; Gullberg, A.; Widegren, B.
J. Mol. Evol. 33, 556-568, 1991
A:Title: The complete nucleotide sequence of the mitochondrial DNA of the fin whale, Balaenoptera physalus
A:Reference number: A58850; MUID:92139449; PMID:1779436
A:Accession: ES5851
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:P24950; UNIPARC:UPI000012894C; GB:X61145; NID:G12772; PIDN:C12772
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; mitochondrion

F:11-209/Domain: cytochrome b homology <CBH>
F:36-52/Domain: cytochrome b6 homology <CB6>
F:81-99/Domain: transmembrane #status predicted <TM1>
F:117-133/Domain: transmembrane #status predicted <TM2>
F:142-146/Region: ubiquinone binding #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:228-245/Domain: transmembrane #status predicted <TM5>
F:323-343/Domain: transmembrane #status predicted <TM6>
F:353-369/Domain: transmembrane #status predicted <TM7>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 534
T11375
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - donkey mitochondrion
C:Species: mitochondrion Equus asinus (donkey)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11375
R:Xu, X.; Gullberg, A.; Arnason, U.
J. Mol. Evol. 43, 438-463, 1996
A:Title: The complete mitochondrial (mtDNA) of the donkey and mtDNA comparisons among felids
A:Reference number: Z17265; MUID:97032591; PMID:8875857
A:Accession: T11375
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <XUX>
A:Cross-references: UNIPROT:P92487; UNIPARC:UPI00001289C6; EMBL:X97337; NID:G1805746; P1:G1805746
C:Genetics:
A:Genome: mitochondrion
A:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation

F:11-209/Domain: cytochrome b homology <CB6>
F:36-52/Domain: cytochrome b6 homology <CB6>
F:81-99/Domain: transmembrane #status predicted <TM1>
F:117-133/Domain: transmembrane #status predicted <TM2>
F:142-146/Region: ubiquinone binding #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:228-245/Domain: transmembrane #status predicted <TM5>
F:323-343/Domain: transmembrane #status predicted <TM6>
F:353-369/Domain: transmembrane #status predicted <TM7>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 288 LGGVLA 293

RESULT 535

T11869
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - horse mitochondrion
C:Species: mitochondrion Equus caballus (domestic horse)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11869
R:Yu, X.; Arnason, U.
Gene 148, 357-362, 1994
A:Title: The complete mitochondrial DNA sequence of the horse, Equus caballus: Extensive
A:Reference number: Z17369; MUID:95047450; PMID:7958969
A:Accession: T11869
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <XUX>
A:Cross-references: UNIPROT:P48665; UNIPARC:UPI00001289F7; EMBL:X79547; NID:9577571; PFI
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:11-339/Domain: cytochrome b homology <CYB>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

|||||

DB 288 LGGVLA 293

RESULT 536

T11259
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - greater Indian rhinoceros
C:Species: mitochondrion Rhinoceros unicornis (greater Indian rhinoceros)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11259
R:Yu, X.; Janke, A.; Arnason, U.
Mol. Biol. Evol. 13, 1167-1173, 1996
A:Title: The complete mitochondrial DNA sequence of the greater Indian rhinoceros, Rhinoceros unicornis
A:Reference number: Z17256; MUID:97051708; PMID:8896369
A:Accession: T11259
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <XUX>
A:Cross-references: UNIPROT:Q96071; UNIPARC:UPI0000128ACB; EMBL:X97336; NID:g1666193; PFI
A:Experimental source: kidney
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:11-339/Domain: cytochrome b homology <CYB>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

|||||

DB 288 LGGVLA 293

RESULT 537

T11349
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - aardvark mitochondrion
C:Species: mitochondrion Orycteropus afer (aardvark)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11349
R:Arnason, U.; Gullberg, A.; Janke, A.
Proc. R. Soc. Lond. B Biol. Sci. 266, 339-345, 1999
A:Title: The mitochondrial DNA molecule of the aardvark, Orycteropus afer, and the posi
A:Reference number: Z17263; MUID:99197468; PMID:10097395
A:Accession: T11349
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q9THD6; UNIPARC:UPI0000090DS2; EMBL:Y18475; NID:g4691353; PFI
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
A>Note: cytb
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:11-339/Domain: cytochrome b homology <CYB>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - aardvark mitochondrion
C:Species: mitochondrion Orycteropus afer (aardvark)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11349
R:Arnason, U.; Gullberg, A.; Janke, A.
Proc. R. Soc. Lond. B Biol. Sci. 266, 339-345, 1999
A:Title: The mitochondrial DNA molecule of the aardvark, Orycteropus afer, and the posi
A:Reference number: Z17263; MUID:99197468; PMID:10097395
A:Accession: T11349
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q9THD6; UNIPARC:UPI0000090DS2; EMBL:Y18475; NID:g4691353; PFI
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
A>Note: cytb
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:11-339/Domain: cytochrome b homology <CYB>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

|||||

DB 288 LGGVLA 293

RESULT 538

T11492
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rabbit mitochondrion
C:Species: mitochondrion Oryctolagus cuniculus (domestic rabbit)
C:Date: 16-Jul-1999 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C:Accession: T11492; S13823
R:Giusti, C.; Gullberg, A.; Arnason, U.
Genomics 50, 161-169, 1998
A:Title: The complete mitochondrial DNA sequence of the rabbit, Oryctolagus cuniculus.
A:Reference number: Z17275; MUID:98317530; PMID:9653643
A:Accession: T11492
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <GIS>
A:Cross-references: UNIPROT:P34863; UNIPARC:UPI0000128AC6; EMBL:AJ001588; NID:g3293006;
A>Note: the termination resulting from transcript polyadenylation is shown
R:Mignotte, F.; Gueride, M.; Champagne, A.M.; Mounolou, J.C.
Eur. J. Biochem. 194, 561-571, 1990
A:Title: Direct repeats in the non-coding region of rabbit mitochondrial DNA. Involvement
A:Reference number: S13823; MUID:91099332; PMID:2269281
A:Accession: S13823
A:Molecule type: DNA
A:Residues: 211-379 <EUR>
A:Cross-references: UNIPARC:UPI000016D6B1; EMBL:X54172; NID:g14141; PIDN:CAA38105.1; PFI
C:Genetics:
A:Gene: cytb
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:11-339/Domain: cytochrome b homology <CYB>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 539

S58462
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Steller sea lion mitochondrion
C;Species: mitochondrion Eumetopias jubatus (Steller sea lion)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S58462
R;Arnason, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A;Title: A molecular view of pinned relationships with particular emphasis on the true
A;Reference number: S58447; MUID:95230701; PMID:7714914
A;Accession: S58462
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Cross-references: UNIPROT:Q34471; UNIPARC:UPI000008FBB9; EMBL:X82311; NID:G693959; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;142-146/Region: ubiquinone binding #status predicted
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;223-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 540

T10998
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b [similarity] - Chrysomys pictus
C;Species: mitochondrion Chrysomys pictus
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T10998
R;Sorenson, M.D.; Dimcheff, D.E.; Ast, J.C.; Yuri, T.; Mindell, D.P.
submitted to the EMBL Data Library, September 1998
A;Description: Complete mitochondrial DNA sequences for five birds and a turtle.
A;Reference number: Z17239
A;Accession: T10998
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Cross-references: UNIPROT:Q9XXK7; UNIPARC:UPI000008C439; EMBL:AF069423; NID:G4530187;
A;Residues: 1-379 <SOR>
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F;84.183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98.197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 289 LGGVLA 294

RESULT 541

D90625
cytochrome b [imported] - Casuarium casuarium mitochondrion
C;Species: mitochondrion Casuarium casuarium
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 09-Jul-2004
C;Accession: D90625
R;Hadrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A;Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogeny
A;Reference number: A99613; MUID:21263106; PMID:11370967
A;Accession: D90625
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-379 <KUR>
A;Cross-references: UNIPROT:Q957Y9; UNIPARC:UPI00000914CB; GB:NC_002778; NID:G14141869;
C;Genetics:
A;Gene: CYTB
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: heme; iron; metalloprotein; mitochondrion
F;84.183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98.197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 289 LGGVLA 294

RESULT 542

D90615
cytochrome b [imported] - Dromaius novaehollandiae mitochondrion
C;Species: mitochondrion Dromaius novaehollandiae
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 09-Jul-2004
C;Accession: D90615
R;Hadrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A;Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogeny
A;Reference number: A99613; MUID:21263106; PMID:11370967
A;Accession: D90615
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-379 <KUR>
A;Cross-references: UNIPROT:Q957Z9; UNIPARC:UPI0000095F8A; GB:NC_002784; NID:G14141883;
C;Genetics:
A;Gene: CYTB
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: heme; iron; metalloprotein; mitochondrion
F;84.183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98.197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||

Db 289 LGGVLA 294

RESULT 543

Tl1178
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - greater rhea mitochondrion
C:Species: mitochondrion Rhea americana (greater rhea, common rhea)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: Tl1178; Tl1426
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242
A:Accession: Tl1178
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <MIN>
A:CROSS-references: UNIPROT:Q9TGW6; UNIPARC:UPI0000096EF2; EMBL:AF090339; NID:G4894475;
R:Harlid, A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z17270
A:Accession: Tl1426
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-29, 'T', 31-379 <HAR>
A:CROSS-references: UNIPARC:UPI0000098AD5; EMBL:Y16884; PIDN:CAA76511.2
C:Genetics:
A:Genome: mitochondrion
A:Gene: cytb
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein, electron transfer; heme; iron; metalloprotein; mitochondrion;
F:12-340/Domain: cytochrome b homology <CYB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 289 LGGVLA 294

RESULT 544

Tl1530
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - ostrich mitochondrion
C:Species: mitochondrion Struthio camelus (ostrich)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: Tl1530
R:Harlid, A.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z17278
A:Accession: Tl1530
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <HAR>
A:CROSS-references: UNIPROT:Q03548; UNIPARC:UPI0000128B13; EMBL:Y12025; PIDN:CAA72755.1
A:Experimental source: ostrich
C:Genetics:
A:Gene: cytb
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:12-340/Domain: cytochrome b homology <CYB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 289 LGGVLA 294

RESULT 545

D90613
cytochrome b [imported] - Struthio camelus mitochondrion
C:Species: mitochondrion Struthio camelus
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 09-Jul-2004
C:Accession: D90613
R:Haddrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogen
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: D90613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <KUR>
A:CROSS-references: UNIPROT:Q03548; UNIPARC:UPI0000128B13; GB:NC_002785; NID:G14141811;
C:Genetics:
A:Gene: CYTB
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: heme; iron; metalloprotein; mitochondrion
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 289 LGGVLA 294

RESULT 546

D90627
cytochrome b [imported] - Eudromia elegans mitochondrion
C:Species: mitochondrion Eudromia elegans
C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 09-Jul-2004
C:Accession: D90627
R:Haddrath, O.; Baker, A.J.
Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: D90627
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <KUR>
A:CROSS-references: UNIPROT:Q958A4; UNIPARC:UPI0000095248; GB:NC_002772; NID:G14141814;
C:Genetics:
A:Gene: CYTB
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: heme; iron; metalloprotein; mitochondrion
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 289 LGGVLA 294

RESULT 547

D90619
cytochrome b [imported] - Apteryx haastii mitochondrion
C:Species: mitochondrion Apteryx haastii
C>Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 09-Jul-2004
C:Accession: D90619
R:Hadrath, O.; Baker, A.J.

Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001
A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogene
A:Reference number: A99613; MUID:21263106; PMID:11370967
A:Accession: D90619

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <KUR>
A:Cross-references: UNIPROT:Q958C0; UNIPARC:UPI0000092AE9; GB:NC_002782; NID:gl14141939;

A:Gene: CYTB
C:Genetics:
A:Genome: mitochondrion

A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: heme; iron; metalloprotein; mitochondrion

F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted
Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22

Db 289 LGGVLA 294

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Phacochoerus africanus mit
C:Species: mitochondrion Phacochoerus africanus
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58057
R:Randi, E.; Lucchini, V.; Diong, C.
submitted to the EMBL Data Library, July 1995
A:Description: Evolutionary Genetics of the suiformes.

A:Reference number: S58021
A:Accession: S58057
A:Molecule type: DNA

A:Residues: 1-379 <RAN>
A:Cross-references: UNIPROT:Q36471; UNIPARC:UPI0000090037; EMBL:Z50090; NID:g902686; PID:
A:Experimental source: liver

C:Genetics:
A:Genome: mitochondrion

A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;

F:11-209/Domain: cytochrome b homology <CBH>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>

F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted
F:178-200/Domain: transmembrane #status predicted <TM4>

F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>

F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22

Db 288 LGGVLA 293

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 550

S58085
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Babyrousa babyrussa mitoch
C:Species: mitochondrion Babyrousa babyrussa
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58085

R:Randi, E.; Lucchini, V.; Diong, C.
submitted to the EMBL Data Library, July 1995
A:Description: Evolutionary Genetics of the suiformes.

A:Reference number: S58021
A:Accession: S58085
A:Molecule type: DNA

A:Residues: 1-379 <RAN>
A:Cross-references: UNIPROT:Q36324; UNIPARC:UPI00000991A9; EMBL:Z50106; NID:g908866; PID:

A:Experimental source: liver
C:Genetics:
A:Genome: mitochondrion

A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;

F:11-209/Domain: cytochrome b homology <CBH>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>

F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted
F:178-200/Domain: transmembrane #status predicted <TM4>

F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>

F:323-343/Domain: transmembrane #status predicted <TM7>

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22

Db 288 LGGVLA 293

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 549

S58057

F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 551
S58456
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - polar bear mitochondrion
C:Species: mitochondrion Ursus maritimus (polar bear)
C:Date: 19-Mar-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: S58456
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58456
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q36082; UNIPARC:UPI000016D681; EMBL:X82309; NID:9693987; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: transmembrane #status predicted <TM1>
F:36-52/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 552
S58459
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - New Zealand fur seal mitochondrion
C:Species: mitochondrion Arctocepalus forsteri (New Zealand fur seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58459
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58459
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q33697; UNIPARC:UPI0000128921; EMBL:X82293; NID:9693953; PFI

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: transmembrane #status predicted <TM1>
F:36-52/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 553
S58464
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - leopard seal mitochondrion
C:Species: mitochondrion Hydrurga leptonyx (leopard seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58464
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58464
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q34732; UNIPARC:UPI0000095PC7; EMBL:X82297; NID:9693963; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: transmembrane #status predicted <TM1>
F:36-52/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293


```

Query Match      5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred.No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 22
Db      288 LGGVLA 293
|||||

RESULT 554
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Antarctic fur seal mitochon
C:Species: mitochondrion Arctocephalus gazella (Antarctic fur seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58458
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinned relationships with particular emphasis on the true
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58458
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q33688; UNIPARC:UPI0000128922; EMBL:X82292; NID:G693951; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CB6>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:136-52/Domain: transmembrane #status predicted <TM1>
F:117-133/Domain: transmembrane #status predicted <TM2>
F:142-146/Region: ubiquinone binding #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match      5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred.No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 22
Db      288 LGGVLA 293
|||||

RESULT 555
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Spermophilus richardsonii
C:Species: mitochondrion Spermophilus richardsonii
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: A53077
R:Thomas, W.K.; Martin, S.L.
Mol. Phylogenet. Evol. 2, 330-336, 1993
A:Title: A recent origin of marmots.
A:Reference number: A53077; MUID:94326000; PMID:8049781
A:Accession: A53077
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-379 <RES>
A:Cross-references: UNIPROT:P49341; UNIPARC:UPI0000128B0A; GB:S73150; NID:G639944; PIDN:
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CB6>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match      5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred.No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 22
Db      288 LGGVLA 293
|||||

RESULT 556
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - ribbon seal mitochondrion
C:Species: mitochondrion Phoca fasciata (ribbon seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58447
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinned relationships with particular emphasis on the true
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58447
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q35438; UNIPARC:UPI0000128A8B; EMBL:X82302; NID:G693969; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:136-52/Domain: transmembrane #status predicted <TM1>
F:117-133/Domain: transmembrane #status predicted <TM2>
F:142-146/Region: ubiquinone binding #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match      5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred.No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 22
Db      288 LGGVLA 293
|||||

RESULT 557
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - harp seal mitochondrion
C:Species: mitochondrion Phoca groenlandica (harp seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58448
R:Arnasson, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinned relationships with particular emphasis on the true
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58448
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q35457; UNIPARC:UPI0000128A8C; EMBL:X82303; NID:G693971; PID
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
```


C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b - spotted seal mitochondrion; C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion; F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:288-304/Domain: transmembrane #status predicted <TM7>
F:323-343/Domain: transmembrane #status predicted <TM8>
F:353-369/Domain: transmembrane #status predicted <TM9>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 558
S58450
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - spotted seal mitochondrion
C:Species: Mitochondrion Phoca largha (spotted seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58450
R:Arnsaon, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58450
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q35505; UNIPARC:UPI0000128A8B; EMBL:X82305; NID:g693975; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b - spotted seal mitochondrion; C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion; F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:288-304/Domain: transmembrane #status predicted <TM7>
F:323-343/Domain: transmembrane #status predicted <TM8>
F:353-369/Domain: transmembrane #status predicted <TM9>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 559
S58460
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - spotted seal mitochondrion
C:Species: Mitochondrion Phoca largha (spotted seal)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58460
R:Arnsaon, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58460
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q35505; UNIPARC:UPI0000128A8B; EMBL:X82305; NID:g693975; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b - spotted seal mitochondrion; C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion; F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:288-304/Domain: transmembrane #status predicted <TM7>
F:323-343/Domain: transmembrane #status predicted <TM8>
F:353-369/Domain: transmembrane #status predicted <TM9>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 559
S58460

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Cystophora cristata mitoc
C:Species: mitochondrion Cystophora cristata
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S58460
R:Arnsaon, U.; Bodin, K.; Gullberg, A.; Ledje, C.; Mouchaty, S.
J. Mol. Evol. 40, 78-85, 1995
A:Title: A molecular view of pinniped relationships with particular emphasis on the true seal
A:Reference number: S58447; MUID:95230701; PMID:7714914
A:Accession: S58460
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <ARN>
A:Cross-references: UNIPROT:Q34070; UNIPARC:UPI0000128993; EMBL:X82294; NID:g693955; PFI
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b - spotted seal mitochondrion; C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion; F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:142-146/Region: ubiquinone binding #status predicted <TM4>
F:178-200/Domain: transmembrane #status predicted <TM5>
F:221-339/Domain: plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM6>
F:288-304/Domain: transmembrane #status predicted <TM7>
F:323-343/Domain: transmembrane #status predicted <TM8>
F:353-369/Domain: transmembrane #status predicted <TM9>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 560
C84577
probable nucleosome assembly protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84577
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, L.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84577
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <STO>
A:Cross-references: UNIPROT:Q9ZUP3; UNIPARC:UPI00000AC55A; GB:AE002093; NID:g4191778; PFI
C:Genetics:
A:Gene: At2g19480
A:Map position: 2

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYDEME 63
|||||
Db 64 QYDEME 69

RESULT 561

I48134
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - American spiny rat (Echimys
C;Species: mitochondrion Echimys chrysurus (American spiny rat)
C;Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 31-Dec-2004
C;Accession: I48134
R;da Silva, M.N.F.; Patton, J.L.
Mol. Phylogenet. Evol. 2, 243-255, 1993
A;Title: Amazonian phylogeography: mtDNA sequence variation in arboreal echimyid rodents
A;Reference number: A49605; MUID:94184505; PMID:8136924
A;Accession: I48134
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-379 <RES>
A;Cross-references: UNIPROT:Q34420; UNIPARC:UPI0000096931; GB:L23341; NID:9995848; PIDN:
C;Genetics:
A;Genetic code: SGC1
C;Superfamily: cytochrome b homology; cytochrome b6 homology; plastoquinol-plastocyanin
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-339/Domain: cytochrome b6 homology <CB6>
F;11-339/Domain: cytochrome b6 homology <CB6>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;83-182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 379;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22

Db 288 LGGVLA 293

RESULT 562

CBHU
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - human mitochondrion
C;Species: mitochondrion Homo sapiens (man)
C;Date: 22-May-1981 #sequence_revision 23-Oct-1981 #text_change 03-Jun-2002
C;Accession: A00151; I57452
R;Anderson, S.; Bankier, A.T.; Barrell, B.G.; de Bruijn, M.H.L.; Coulson, A.R.; Drouin,
Nature 290, 457-465, 1981
A;Title: Sequence and organization of the human mitochondrial genome.
A;Reference number: A00151; MUID:81173052; PMID:7219534
A;Accession: A00151
A;Molecule type: DNA
A;Residues: 1-380 <AND>
A;Cross-references: UNIPARC:UPI00001289F8; EMBL:V00662; NID:G13003; PIDN:CAA24038.1; PID
R;Spurr, N.K.; Bodmer, W.P.
Mol. Biol. Med. 2, 239-249, 1984
A;Title: Serendipitous cloning of a mitochondrial cDNA and its polymorphism.
A;Reference number: I57452; MUID:86064879; PMID:6100559
A;Accession: I57452
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 269-380 <GPU>
A;Cross-references: UNIPARC:UPI000016D57B; GB:M28016; NID:G337203; PIDN:AAA31851.1; PID:
C;Genetics:
A;Gene: GDB:MTCYB
A;Map position: MTH14747-15887
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Complex: the transmembrane complex includes cytochrome b, cytochrome c1 (see PIR:S0068
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>

F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22

Db 288 LGGVLA 293

RESULT 563

CBXL

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - African clawed frog mitoch
C;Species: mitochondrion Xenopus laevis (African clawed frog)
C;Date: 28-Feb-1986 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C;Accession: A23955; A00155
R;Dunon-Bluteau, D.; Volovitch, M.; Brun, G.
Gene 36, 65-78, 1985
A;Title: Nucleotide sequence of a Xenopus laevis mitochondrial DNA fragment containing t
A;Reference number: A23955; MUID:86056961; PMID:2415430
A;Accession: A23955
A;Molecule type: DNA
A;Residues: 1-380 <DUN>
A;Cross-references: UNIPROT:P00160; UNIPARC:UPI0000171CE8; GB:M10188
A;Note: the authors state that this sequence corrects that which was reported in referen
R;Jee, B.A.; Ma, D.P.; Wilson, R.K.; Wong, J.F.H.
J. Biol. Chem. 260, 9759-9774, 1985
A;Title: The complete nucleotide sequence of the Xenopus laevis mitochondrial genome.
A;Reference number: A00155; MUID:85261388; PMID:4019494
A;Accession: A00155
A;Molecule type: DNA
A;Residues: 1-71, 'F', 73-77, 'L', 79-87, 'L', 89-153, 'K', 155-164, 'SL', 167-284, 'M', 287-333, 'L',
A;Cross-references: UNIPARC:UPI0000163901; GB:M10217; GB:X02890; NID:G343717; PIDN:AAA66
C;Genetics:
A;Gene: cob
A;Genome: mitochondrion
A;Genetic code: SGC1
C;Function:
A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;12-340/Domain: cytochrome b6 homology <CBH>
F;12-210/Domain: cytochrome b6 homology <CB6>
F;37-53/Domain: transmembrane #status predicted <TM1>
F;82-100/Domain: transmembrane #status predicted <TM2>
F;118-134/Domain: transmembrane #status predicted <TM3>
F;179-201/Domain: transmembrane #status predicted <TM4>
F;222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F;230-246/Domain: transmembrane #status predicted <TM5>
F;289-305/Domain: transmembrane #status predicted <TM6>
F;324-344/Domain: transmembrane #status predicted <TM7>
F;354-370/Domain: transmembrane #status predicted <TM8>
F;84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 Db 289 LGGVLA 294

RESULT 564
 S04840
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b [similarity] - white sturgeon
 C:Species: mitochondrion Acipenser transmontanus (white sturgeon)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
 C:Accession: S04840
 R:Brown, J.R.; Gilbert, T.L.; Kowbel, D.J.; O'Hara, P.J.; Buroker, N.E.; Beckenbach, A.T
 Nucleic Acids Res. 17, 4389, 1989
 A:Title: Nucleotide sequence of the apocytochrome B gene in white sturgeon mitochondrial
 A:Reference number: S04840; MUID:89296501; PMID:2740232
 A:Accession: S04840
 A:Molecule type: DNA
 A:Residues: 1-380 <BRO>
 A:Cross-references: UNIPROT:P11669; UNIPARC:UPI00001288EC; EMBL:X14944; NID:g12749; PIDN
 A:Note: the termination resulting from transcript polyadenylation is shown
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 Db 288 LGGVLA 293

RESULT 565
 S10198
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - chicken mitochondrion
 C:Species: mitochondrion Gallus gallus (chicken)
 C:Date: 31-Dec-1990 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
 C:Accession: S10198
 R:Desjardins, P.; Morais, R.
 J. Mol. Biol. 212, 599-634, 1990
 A:Title: Sequence and gene organization of the chicken mitochondrial genome. A novel gen
 A:Reference number: S10187; MUID:90230301; PMID:2329578
 A:Accession: S10198
 A:Molecule type: DNA
 A:Residues: 1-380 <DES>
 A:Cross-references: UNIPROT:P18946; UNIPARC:UPI0000000244; EMBL:X52392; NID:g12960; PIDN
 C:Genetics:
 A:Gene: cytB
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:12-340/Domain: cytochrome b homology <CBH>

F:12-210/Domain: cytochrome b6 homology <CB6>
 F:37-53/Domain: transmembrane #status predicted <TM1>
 F:82-100/Domain: transmembrane #status predicted <TM2>
 F:118-134/Domain: transmembrane #status predicted <TM3>
 F:179-201/Domain: transmembrane #status predicted <TM4>
 F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:230-246/Domain: transmembrane #status predicted <TM5>
 F:289-305/Domain: transmembrane #status predicted <TM6>
 F:324-344/Domain: transmembrane #status predicted <TM7>
 F:354-370/Domain: transmembrane #status predicted <TM8>
 F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 Db 289 LGGVLA 294

RESULT 566
 S36011
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - common carp mitochondrion
 C:Species: mitochondrion Cyprinus carpio (common carp)
 C:Date: 31-Dec-1993 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
 C:Accession: S36011; E44651
 R:Chang, Y.S.; Huang, F.L.
 submitted to the EMBL Data Library, July 1991
 A:Description: The cDNA and primary structure of pregrowth hormones of three species of
 A:Reference number: S21910
 A:Accession: S36011
 A:Molecule type: DNA
 A:Residues: 1-380, 'C', <CHAI>
 A:Cross-references: UNIPROT:P24951; UNIPARC:UPI0000128992; EMBL:X61010; NID:g436882
 A:Note: GenBank entry M10000 P100000171CE9; MUID:94223691; PMID:8169959
 R:Chang, Y.S.; Huang, F.L.; Lo, T.B.
 J. Mol. Evol. 38, 138-155, 1994
 A:Title: The complete nucleotide sequence and gene organization of carp (Cyprinus carpio)
 A:Reference number: A44650; MUID:94223691; PMID:8169959
 A:Accession: E44651
 A:Molecule type: DNA
 A:Residues: 1-301, 'I', 303-380 <CHAI>
 A:Cross-references: UNIPARC:UPI0000171CE9; EMBL:X61010; NID:g436882
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Function:
 A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
 ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
 A:Pathway: oxidative phosphorylation; respiratory chain
 C:Superfamily: cytochrome b; cytochrome b6 homology; cytochrome b6 homology; plastoquinol
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
 F:11-339/Domain: cytochrome b homology <CBH>
 F:11-209/Domain: cytochrome b6 homology <CB6>
 F:36-52/Domain: transmembrane #status predicted <TM1>
 F:81-99/Domain: transmembrane #status predicted <TM2>
 F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 380;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
| | | | |
Db 288 LGGVLA 293

RESULT 567

D59154
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - western lowland gorilla mi
C:Species: mitochondrion Gorilla gorilla gorilla (western lowland gorilla)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Jun-2002
C:Accession: D59154
R:Yu, X.; Arnason, U.
Mol. Biol. Evol. 13, 691-698, 1996
A:Title: A complete sequence of the mitochondrial genome of the Western lowland gorilla.
A:Reference number: Z17269; MUID:96212991; PMID:8676744
A:Accession: D59154
A>Status: preliminary; nucleic acid sequence not shown; translation not shown; translate
A:Molecule type: DNA
A:Residues: 1-380 <XU>
A:Cross-references: UNIPARC:UPI0000174C8C; GB:X93347; NID:gl304307; GSPDB:GN00106
A:Note: submitted to GenBank, November 1995
A:Note: this translation is not annotated in GenBank entry GCMITG, release 114.0
C:Genetics:
A:Gene: cytb
A:Genome: mitochondrion
A:Genetic code: SGCl
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CVB>
F:14-209/Domain: cytochrome b6 homology <CB6>
F:12-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:183.182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:197.196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
| | | | |
Db 288 LGGVLA 293

RESULT 568

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - redhead mitochondrion
C:Species: mitochondrion Aythya americana (redhead)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: T11033
R:Mindell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A:Title: Multiple independent origins of mitochondrial gene order in birds.
A:Reference number: Z17242
A:Accession: T11033
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-380 <JOH>
A:Cross-references: UNIPROT:O63748; UNIPARC:UPI0000094BB7; EMBL:AF090337; NID:G4887659;
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCl
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:12-340/Domain: cytochrome b homology <CVB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:184.183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:198.197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
| | | | |
Db 289 LGGVLA 294

RESULT 569

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - turkey mitochondrion
C:Species: mitochondrion Meleagris gallopavo (common turkey)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: I51374
R:Kornegay, J.R.; Kocher, T.D.; Williams, L.A.; Wilson, A.C.
J. Mol. Evol. 37, 367-379, 1993
A:Title: Pathways of lysozyme evolution inferred from the sequences of cytochrome b in b
A:Reference number: I50092; MUID:94141938; PMID:8308906
A:Accession: I51374
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-380 <KOR>
A:Cross-references: UNIPROT:P50663; UNIPARC:UPI0000128A3B; GB:L08381; NID:G343475; PIDN:
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCl
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:12-340/Domain: cytochrome b homology <CB6>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:84.183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:198.197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
| | | | |
Db 289 LGGVLA 294

RESULT 570

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - coelacanth mitochondrion
C:Species: mitochondrion Latimeria chalumnae (coelacanth)
C:Date: 30-Oct-1998 #sequence_revision 30-Oct-1998 #text_change 09-Jul-2004
C:Accession: E58893
R:Zardoya, R.; Meyer, A.
Genetics 146, 995-1010, 1997
A:Title: The complete DNA sequence of the mitochondrial genome of a "living fossil," the
A:Reference number: A58892; MUID:97358858; PMID:9215903
A:Accession: E58893
A>Status: preliminary; nucleic acid sequence not shown; translation not shown; not compa
A:Molecule type: DNA
A:Residues: 1-380 <ZAR>
A:Cross-references: UNIPROT:O03176; UNIPARC:UPI0000128A11; GB:U82228; NID:G1916917; PID:
A:Note: submitted to GenBank/EMBL/DBJ December, 1996
C:Genetics:
A:Gene: Cytb
A:Map position: FOR14343-15485
A:Genome: mitochondrion
A:Genetic code: SGCl
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:229-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>

F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 571
Tl1114
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Florometra serratissima m
C:Species: mitochondrion Florometra serratissima
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: Tl1114
R:Scouras, A.; Smith, M.J.
submitted to the EMBL Data Library, February 1998
A:Description: The complete mitochondrial genome of the crinoid Florometra serratissima.
A:Reference number: Z17249
A:Accession: Tl1114
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-380 <SCO>
A:Cross-references: UNIPROT:O63585; UNIPARC:UPI0000090197; EMBL:AF049132; NID:g2970420;
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:12-340/Domain: cytochrome b homology <CYB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 289 LGGVLA 294

RESULT 572
Tl1518
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b [similarity] - baboon mitoch
C:Species: mitochondrion Papio hamadryas (baboon)
C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: Tl1518
R:Arnaason, U.; Gullberg, A.; Janke, A.
J. Mol. Evol. 47, 718-727, 1998
A:Title: Molecular timing of primate divergences as estimated by two non-primate calibra
A:Reference number: Z17277; MUID:99065765; PMID:9847414
A:Accession: Tl1518
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-380 <ARN>
A:Cross-references: UNIPROT:Q9T4X4; UNIPARC:UPI0000094469; EMBL:Y18001; NID:g4049475; PI
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 573
Tl1845
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - common gibbon mitochondri
C:Species: mitochondrion Hylobates lar (common gibbon, white-handed gibbon)
C>Date: 16-Jul-1999 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C:Accession: Tl1845
R:Arnaason, U.; Gullberg, A.; Xu, X.
Hereditas 124, 185-189, 1996
A:Title: A complete mitochondrial DNA molecule of the white-handed gibbon, Hylobates la
A:Reference number: Z17353
A:Accession: Tl1845
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-380 <ARN>
A:Cross-references: UNIPROT:O95711; UNIPARC:UPI00001289FC; EMBL:X99256; PIDN:CAA67640.1
A:Experimental source: isolate Ester
A>Note: the termination resulting from transcript polyadenylation is shown
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGCI
A:Note: cytb
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative pho
F:11-339/Domain: cytochrome b homology <CYB>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 574
Tl1335
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Corvus frugilegus mitoch
C:Species: mitochondrion Corvus frugilegus
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C:Accession: Tl1335
R:Harlid, A.; Arnaason, U.
Proc. R. Soc. Lond. B Biol. Sci. 266, 305-309, 1999
A:Title: Analyses of mitochondrial DNA nest ratite birds within the Neognathae-supportit
A:Reference number: Z17262
A:Accession: Tl1335
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-380 <HAR>
A:Cross-references: UNIPROT:O79386; UNIPARC:UPI0000091E07; EMBL:Y18522; PIDN:CAA77206.1
C:Genetics:
A:Genome: mitochondrion
A:Note: cytb
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative pho
F:12-340/Domain: cytochrome b homology <CYB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;

[illegible]

RESULT 579

T09869

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - rainbow trout mitochondrion

C:Species: mitochondrion Oncorhynchus mykiss (rainbow trout)

C:Date: 16-Jul-1999 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004

C:Accession: T09869

R:Zardoya, R.; Garrido-Perterra, A.; Bautista, J.M.

J. Mol. Evol. 41, 942-951, 1995

A:Title: The complete nucleotide sequence of the mitochondrial DNA genome of the rainbow

A:Reference number: Z16890; MUID:96139027; PMID:8587139

A:Accession: T09869

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-380 <ZAR>

A:Cross-references: UNIPROT:P48173; UNIPARC:UPI0000128A6D; EMBL:L29771; NID:g1246865; PI

A:Experimental source: liver

A>Note: the termination resulting from transcript polyadenylation is shown

C:Genetics:

A:Gene: Cyt b

A:Genome: mitochondrion

A:Genetic code: SGCI

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;

F:11-339/Domain: cytochrome b6 homology <CYB>

F:11-339/Domain: cytochrome b6 homology <CB6>

F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 288 LGGVLA 293

RESULT 580

T09959

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Atlantic salmon mitochondrion

C:Species: mitochondrion Salmo salar (Atlantic salmon)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C:Accession: T09959

R:Huret, C.D.; Bartlett, S.B.; Bruce, I.J.; Davidson, W.S.

submitted to the EMBL Data Library, October 1998

A:Description: The complete nucleotide sequence of the mitochondrial DNA of the Atlantic

A:Reference number: Z16904

A:Accession: T09959

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-380 <HUR>

A:Cross-references: UNIPROT:Q35925; UNIPARC:UPI0000128AD5; EMBL:U12143; NID:g3775976; PI

A:Experimental source: liver

C:Genetics:

A:Gene: Cytb

A:Genome: mitochondrion

A:Genetic code: SGCI

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos

F:11-339/Domain: cytochrome b6 homology <CYB>

F:11-339/Domain: cytochrome b6 homology <CB6>

F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 288 LGGVLA 293

RESULT 581

D90623

cytochrome b [imported] - Anomalopteryx didiformis mitochondrion

C:Species: mitochondrion Anomalopteryx didiformis

C:Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 09-Jul-2004

C:Accession: D90623

R:Hadrath, O.; Baker, A.J.

Proc. R. Soc. Lond. B Biol. Sci. 268, 939-945, 2001

A:Title: Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogen

A:Reference number: A99613; MUID:21263106; PMID:11370967

A:Accession: D90623

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-380 <KUR>

A:Cross-references: UNIPROT:Q95YX7; UNIPARC:UPI00000952D1; GB:NC_002779; NID:g14141927;

C:Genetics:

A:Gene: CVTB

A:Genome: mitochondrion

A:Genetic code: SGCI

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: heme; iron; metalloprotein; mitochondrion

F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 289 LGGVLA 294

RESULT 582

S42245

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - muscovy duck mitochondrion

C:Species: mitochondrion Cairina moschata (muscovy duck)

C:Date: 12-Dec-1994 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004

C:Accession: S42245

R:Kornegay, J.R.; Kocher, T.D.; Williams, L.A.; Wilson, A.C.

submitted to the EMBL Data Library, December 1992

A:Description: Pathways of lysosome evolution in birds inferred from the sequences of c;

A:Reference number: S42245

A:Accession: S42245

A:Molecule type: DNA

A:Residues: 1-380 <KOR>

A:Cross-references: UNIPROT:Q34160; UNIPARC:UPI00000986B9; EMBL:L08385; NID:g336527; PI

C:Genetics:

A:Gene: mitochondrion

A:Genetic code: SGCI

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion,

F:12-340/Domain: cytochrome b homology <CBH>

F:12-210/Domain: cytochrome b6 homology <CB6>

F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>

F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 289 LGGVLA 294

RESULT 583

T11086

ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Falco peregrinus mitochondrion
C/Species: mitochondrion Falco peregrinus
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11086
R/Mindeall, D.P.; Sorenson, M.D.; Dimcheff, D.E.
Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
A/Title: Multiple independent origins of mitochondrial gene order in birds.
A/Reference number: Z17242
A/Accession: T11086
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-380 <MIN>
A/Cross-references: UNIPROT:Q9T2G1; UNIPARC:UPI00000983DA; EMBL:AF090338; NID:g4894462;
C/Genetics:
A/Genome: mitochondrion
A/Gene: SGC1
A/Gene: SGC1
C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology
C/Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F:12-340/Domain: cytochrome b homology <CYB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: cytochrome b6 homology
F:184,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 289 LGGVLA 294

RESULT 584
T11299
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Raja radiata mitochondrion
C/Species: mitochondrion Raja radiata
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11299
R/Rasmussen, A.S.; Arnason, U.
Proc. Natl. Acad. Sci. U.S.A. 96, 2177-2182, 1999
A/Title: Molecular studies suggest that cartilaginous fishes have an apical position in
A/Reference number: Z17259; MUID:99162577; PMID:10051614
A/Accession: T11299
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-380 <RAS>
A/Cross-references: UNIPROT:Q9ZX33; UNIPARC:UPI000008E524; EMBL:AF106038; NID:g4406269;
C/Genetics:
A/Gene: CYTB
A/Genome: mitochondrion
A/Gene: SGC1
A/Gene: SGC1
C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology; cytochrome b6 homology
C/Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F:12-340/Domain: cytochrome b homology <CYB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: cytochrome b6 homology
F:184,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 289 LGGVLA 294

RESULT 585
H70590
hypochemical protein RV3230C - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C/Accession: H70590
R/Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: H70590
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-380 <COL>
A/Cross-references: UNIPROT:O05875; UNIPARC:UPI000000D5E45; GB:Z95121; GB:AL123456; NID:g
C/Genetics:
A/Experimental source: strain H37RV
A/Gene: RV3230C
C/Superfamily: phthalate dioxygenase reductase; cytochrome-b5 reductase homology; ferredoxin
F:318-369/Domain: ferredoxin [2Fe-2S] homology <FER>

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 217 LAALAA 222

RESULT 586
DB4295
hypochemical protein Vng1409c [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: DB4295
R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Lethauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A/Title: Genome sequence of Halobacterium species NRC-1.
A/Reference number: A84160; MUID:20504483; PMID:11016950
A/Accession: DB4295
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-380 <STO>
A/Cross-references: UNIPROT:Q9HPY9; UNIPARC:UPI00000638A2; GB:AE004437; NID:g10580914; P
C/Genetics:
A/Gene: VNG1409C

Query Match 5.1%; Score 6; DB 2; Length 380;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
DB 127 VLLGGV 132

RESULT 587
CBMS
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - mouse mitochondrion
C/Species: mitochondrion Mus musculus (house mouse)
C/Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C/Accession: A00153
R/Bibb, M.J.; Van Etten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.
Cell 26, 167-180, 1981
A/Title: Sequence and gene organization of mouse mitochondrial DNA.
A/Reference number: A00153; MUID:82137051; PMID:7332926
A/Accession: A00153
A/Molecule type: DNA
A/Residues: 1-381 <BIB>
A/Cross-references: UNIPROT:P00158; UNIPARC:UPI00000008PF; GB:J01420; NID:g342520; PID:g

R,Cao, Y.; Waddell, P.J.; Okada, N.; Hasegawa, M.
Mol. Biol. Evol. 15, 1637-1646, 1998
A>Title: The complete mitochondrial DNA sequence of the shark (*Mustelus manazo*): Evaluation of the complete mitochondrial genome
A:Reference number: Z17338; MUID:99083431; PMID:9866199
A:Accession: T11776
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-381 <AO>
A:Cross-references: UNIPROT:079571; UNIPARC:UPI0000090AE2; EMBL:AB015962; PIDN:BA033047
A:Experimental source: liver
C:Genetics:
A:Genome: mitochondrion
A:Note: cyf-b
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
F:12-340/Domain: cytochrome b6 homology <CB6>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:223-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 289 LGGVLA 294

RESULT 592
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Protopterus dolloi mitoch
C:Species: mitochondrion Protopterus dolloi
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
A:Accession: S68140
A:Zardoya, R.; Meyer, A.
Genetics 142, 1249-1263, 1996
A>Title: The complete nucleotide sequence of the mitochondrial genome of the lungfish (E
A:Reference number: S68128; MUID:96271539; PMID:8846902
A:Accession: S68140
A:Molecule type: DNA
A:Residues: 1-381 <ZAR>
A:Cross-references: UNIPROT:Q35424; UNIPARC:UPI0000091043; EMBL:L42813; NID:gl161203; PI
C:Genetics:
A:Gene: cytb
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:12-340/Domain: cytochrome b homology <CBH>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:37-53/Domain: transmembrane #status predicted <TM1>
F:82-100/Domain: transmembrane #status predicted <TM2>
F:118-134/Domain: transmembrane #status predicted <TM3>
F:179-201/Domain: transmembrane #status predicted <TM4>
F:222-340/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:230-246/Domain: transmembrane #status predicted <TM5>
F:289-305/Domain: transmembrane #status predicted <TM6>
F:324-344/Domain: transmembrane #status predicted <TM7>
F:354-370/Domain: transmembrane #status predicted <TM8>
F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 381;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 289 LGGVLA 294

RESULT 593
S33573
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - short-tailed opossum (Mon
C:Species: mitochondrion Monodelphis domestica
C:Date: 03-Feb-1994 #sequence_revision 20-Aug-1994 #text_change 09-Jul-2004
A:Accession: S33573
R,Ma, D.P.; Zharkikh, A.; Graur, D.; VandeBerg, J.L.; Li, W.H.
J. Mol. Evol. 36, 327-334, 1993
A>Title: Structure and evolution of opossum, guinea pig, and porcine cytochrome b gene
A:Reference number: S33572; MUID:93301932; PMID:8315653
A:Accession: S33573
A:Molecule type: DNA
A:Residues: 1-382 <MAD>
A:Cross-references: UNIPROT:Q04911; UNIPARC:UPI0000128A44; EMBL:X70673; NID:gl14019; PIDN
A:Note: residue 1 and the corresponding nucleotide sequence are not shown
C:Genetics:
A:Gene: cob
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Function:
A:Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A:Pathway: oxidative phosphorylation; respiratory chain
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>
F:117-133/Domain: transmembrane #status predicted <TM3>
F:178-200/Domain: transmembrane #status predicted <TM4>
F:221-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:223-245/Domain: transmembrane #status predicted <TM5>
F:288-304/Domain: transmembrane #status predicted <TM6>
F:323-343/Domain: transmembrane #status predicted <TM7>
F:353-369/Domain: transmembrane #status predicted <TM8>
F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 382;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 594
S47882
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - North American opossum mit
C:Species: mitochondrion Didelphis virginiana, Didelphis marsupialis virginiana (North A
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
A:Accession: S47882; S42822
R,Janke, A.; Feldmaier-Fuchs, G.; Thomas, W.K.; von Haeseler, A.; Paasebo, S.
Genetics 137, 243-256, 1994
A>Title: The marsupial mitochondrial genome and the evolution of placental mammals.
A:Reference number: S47870; MUID:94333786; PMID:8056314
A:Accession: S47882
A:Molecule type: DNA
A:Residues: 1-382 <JAN>
A:Cross-references: UNIPROT:P41303; UNIPARC:UPI00001289A7; EMBL:Z29573; NID:g452251; PID
C:Genetics:
A:Gene: cyb
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex; m
in
F:11-339/Domain: cytochrome b homology <CBH>
F:11-209/Domain: cytochrome b6 homology <CB6>
F:36-52/Domain: transmembrane #status predicted <TM1>
F:81-99/Domain: transmembrane #status predicted <TM2>

F:117-133/Domain: transmembrane #status predicted <TM3>
 F:178-200/Domain: transmembrane #status predicted <TM4>
 F:221-339/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
 F:229-245/Domain: transmembrane #status predicted <TM5>
 F:288-304/Domain: transmembrane #status predicted <TM6>
 F:323-343/Domain: transmembrane #status predicted <TM7>
 F:353-369/Domain: transmembrane #status predicted <TM8>
 F:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 DB 288 LGGVLA 293

RESULT 595
 D58930
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Cyanidioschyzon merolae m
 C:Species: mitochondrion Cyanidioschyzon merolae
 C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 C:Accession: D58930
 R:Ohta, N.; Sato, N.; Kuroiwa, T.
 Nucleic Acids Res. 26, 5190-5198, 1998
 A:Title: Structure and organization of the mitochondrial genome of the unicellular red a
 A:Reference number: A58930; MUID:99030526; PMID:9801118
 A:Accession: D58930
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-382 <ARN>
 A:Cross-references: UNIPROT:Q9Z2R0; UNIPARC:UPI000008DEA4; GB:D89861; NID:g4115781; PIDN
 C:Genetics:
 A:Gene: cytb
 A:Genome: mitochondrion
 C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyaninol
 C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
 F:9-339/Domain: cytochrome b homology <CYB>
 F:9-209/Domain: cytochrome b6 homology <CB6>
 F:221-339/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
 F:81,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
 F:95,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 DB 288 LGGVLA 293

RESULT 596
 JC2571
 cellulase (EC 3.2.1.4) precursor - Streptomyces rochei (strain A2)
 N:Alternate names: endo-1,4-beta-glucanase; endoglucanase
 C:Species: Streptomyces rochei
 C:Date: 13-Jun-1995 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
 C:Accession: JC2571; S34392
 R:Perito, B.; Hanhart, E.; Irdani, T.; Iqbal, M.; McCarthy, A.J.; Mastromei, G.
 Gene 148, 119-124, 1994
 A:Title: Characterization and sequence analysis of a Streptomyces rochei A2 endoglucanase
 A:Reference number: JC2571; MUID:95011642; PMID:7523249
 A:Accession: JC2571
 A:Molecule type: DNA
 A:Residues: 1-382 <PER>
 A:Cross-references: UNIPROT:Q59963; UNIPARC:UPI0000081DAA; EMBL:X73953; NID:g393391; PID
 A:Note: this cellulolytic strain was isolated from the gut of termites
 C:Genetics:
 A:Gene: egIs
 C:Function:

A:Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as c:
 A:Pathway: cellulose degradation
 C:Keywords: glycosidase; hydrolase; polysaccharide degradation
 F:1-37/Domain: signal sequence #status predicted <SIG>
 F:38-382/Product: endoglucanase #status predicted <MAT>
 F:279-380/Domain: bacterial cellulose-binding domain homology <BCB>
 F:280-379/Disulfide bonds: #status predicted

Query Match 5.1%; Score 6; DB 2; Length 382;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 22 LAALAA 27

RESULT 597
 S54213
 flagellar biosynthetic protein flhB - Yersinia enterocolitica
 C:Species: Yersinia enterocolitica
 C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
 C:Accession: S54213
 R:Faucomnier, A.; Allaoui, A.; van Elsen, A.; Cornelis, G.; Bollen, A.
 submitted to the EMBL Data Library, February 1995
 A:Description: Clustering of flagellar genes around invA, the Yersinia enterocolitica i
 A:Reference number: S54213
 A:Accession: S54213
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <PAU>
 A:Cross-references: UNIPROT:Q56886; UNIPARC:UPI000012A95E; EMBL:Z48169; NID:g793891; PT
 C:Genetics:
 A:Start codon: GTG
 A:Superfamily: flagellar biosynthetic protein flhB; flhB carboxyl-terminal homology
 F:274-347/Domain: flhB carboxyl-terminal homology <PCT>

Query Match 5.1%; Score 6; DB 2; Length 383;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVVL 21
 |||||
 DB 110 LGGVVL 115

RESULT 598
 G75431
 probable Na+/H+ antiporter - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: G75431
 R:White, O.; Eilen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 S.; Smith, H.O.; Vamathavan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: G75431
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <WHI>
 A:Cross-references: UNIPROT:Q9RV80; UNIPARC:UPI00000D3DDC; GB:AE001964; GB:AE000513; NID
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1149
 A:Map position: 1
 C:Superfamily: Aquifex aeolicus Na+/H+-exchanging protein napAI

Query Match 5.1%; Score 6; DB 2; Length 383;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 WVLGG 19
|||||
Db 115 WVLGG 120

RESULT 599

S70984
recP protein - Mycobacterium tuberculosis (strain H37RV)
C/Species: Mycobacterium tuberculosis
C/Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 09-Jul-2004
C/Accession: S70984; A70698
R:Salazar, L.; Fathi, H.; de Rossi, E.; Riccardi, G.; Rios, C.; Cole, S.T.; Takiff, H.E.
Mol. Microbiol. 20, 283-293, 1996
A/Title: Organization of the origins of replication of the chromosomes of Mycobacterium smatis.
A/Reference number: S70980; MUID:96310367; PMID:8733228
A/Accession: S70984
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Map position: 4R
A/Molecule type: DNA
A/Residues: 1-385 <SAL>
A/Cross-references: UNIPROT:Q59586; UNIPARC:UPI000016FBBS; EMBL:X92504; NID:g1321901; PI
A/Note: The nucleotide sequence was submitted to the EMBL Data Library, October 1995
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A/Reference number: A70500; MUID:98295987; PMID:9634230
A/Accession: A70698
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-183, 'EV', 186-385 <COL>
A/Cross-references: UNIPARC:UPI0000133562; GB:Z80233; GB:AL123456; NID:g3261645; PIDN:CA
A/Experimental source: strain H37RV
C/Genetics:
A/Gene: recP
A/Start codon: GTG
C/Superfamily: recF protein

Query Match 5.1%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 255 LAALAA 260

RESULT 600

S69587
hypothetical protein YDR532c - yeast (Saccharomyces cerevisiae)
C/Species: Saccharomyces cerevisiae
C/Date: 22-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C/Accession: S69587
R:Dietrich, F.S.
submitted to the EMBL Data Library, August 1995
A/Description: The sequence of S. cerevisiae cosmid 8166, 9787, 9717, and lambda 3073.
A/Reference number: S69553
A/Accession: S69587
A/Molecule type: DNA
A/Residues: 1-385 <DIE>
A/Cross-references: UNIPROT:Q04431; UNIPARC:UPI000006A39B; EMBL:U33057; NID:g927764; PID
C/Genetics:
A/Gene: SGD,KRE28; MIPS:YDR532C
A/Cross-references: SGD:S0002940
A/Map position: 4R

Query Match 5.1%; Score 6; DB 2; Length 385;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 106 TWQKL 111

Db 226 TWQKL 231
|||||

RESULT 601

T11832
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Atlantic cod mitochondrion
C/Species: mitochondrion Gadus morhua (Atlantic cod)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C/Accession: T11832; S45353
R:Johansen, S.; Bakke, I.
Mol. Marine Biol. Biotechnol. 5, 203-214, 1996
A/Title: The complete mitochondrial DNA sequence of Atlantic cod, Gadus morhua: Relevanc
A/Reference number: Z17351; MUID:96414925; PMID:8817926
A/Accession: T11832
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-386 <JOH1>
A/Cross-references: UNIPROT:Q9MJ7; UNIPARC:UPI00000993CA; EMBL:X99772; PIDN:CAA68117.1
A/Experimental source: norwegian coastal stock (NC-1)
R:Johansen, S.; Johansen, T.
Biochim. Biophys. Acta 1218, 213-217, 1994
A/Title: Sequence analysis of 12 structural genes and a novel non-coding region from mit
A/Reference number: S45350; MUID:94289483; PMID:8018725
A/Accession: S45353
A/Molecule type: DNA
A/Residues: 1-380 <JOH2>
A/Cross-references: UNIPARC:UPI00001289DD; EMBL:X76365; NID:9525249; EMBL:X76366; NID:95
A/Experimental source: strain NC1
A/Note: strain AN1 has also been sequenced by the authors
C/Genetics:
A/Gene: cytb
A/Genome: cytb
A/Genetic code: SGC1
C/Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C/Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phos
P:11-339/Domain: cytochrome b homology <CBH>
P:11-209/Domain: cytochrome b6 homology <CB6>
P:121-339/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
P:83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
P:97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicte

Query Match 5.1%; Score 6; DB 2; Length 386;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 602

T11286
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b [similarity] - American alli
C/Species: mitochondrion Alligator mississippiensis (American alligator)
C/Date: 16-Jul-1999 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C/Accession: T11286; T43696
R:Janke, A.
submitted to the EMBL Data Library, July 1998
A/Description: The complete mitochondrial genome of Alligator mississippiensis and the s
A/Reference number: Z17258
A/Accession: T11286
A/Molecule type: DNA
A/Residues: 1-386 <JAN>
A/Cross-references: UNIPROT:O47878; UNIPARC:UPI000005D8BA; EMBL:Y13113; PIDN:CAA73573.1.
A/Experimental source: adult; liver
A/Note: the termination resulting from transcript polyadenylation is shown
R:Sorenson, M.D.; Dimcheff, D.E.; Ast, J.C.; Yuri, T.; Mindell, D.P.
submitted to the EMBL Data Library, June 1998
A/Description: Primers for a PCR-based approach to complete mitochondrial genome sequenc
A/Reference number: Z22631
A/Accession: T43696

A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-386 <SOR>
A:Cross-references: UNIPARC:UPI000005D8BA; EMBL:AF069428; NID:g3219514; PIDN:AAD09991.1;
C:Genetics:
A:Genome: mitochondrion
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; cytochrome b6 homology; plastocyanin
C:Keywords: electron transfer; heme; iron; metalloprotein; mitochondrion; oxidative phosphorylation
F:12-340/Domain: cytochrome b6 homology <CVB>
F:12-210/Domain: cytochrome b6 homology <CB6>
F:222-340/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>
F:84_183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:98_197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 386;
Best Local Similarity 100.0%; Pred.No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
| | | | |
Db 289 LGGVLA 294

RESULT 603
AD0218
Flagellar biosynthetic protein FlhB [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD0218
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
delo-Tarraga, A.M.; Chillingworth, T.; Cronan, A.; Davies, R.M.; Davis, P.; Dougan, G.;
lin, M.; Rutherford, K.; Simmonds, T.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0218
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-386 <KUR>
A:Cross-references: UNIPROT:Q8ZFEC4; UNIPARC:UPI00000DC8CC; GB:AL590842; PIDN:CAC90608.1;
C:Genetics:
A:Gene: flhB
C:Superfamily: flagellar biosynthetic protein flhB; flhB carboxyl-terminal homology

Query Match 5.1%; Score 6; DB 2; Length 386;
Best Local Similarity 100.0%; Pred.No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGVL 21
| | | | |
Db 110 LLGVL 115

RESULT 604
CBASN
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Emericella nidulans mitoch
C:Species: mitochondrion Emericella nidulans, Aspergillus nidulans
C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 09-Jul-2004
C:Accession: AA0157
R:Waring, R.B.; Davies, R.W.; Lee, S.; Grisi, E.; Berks, M.M.; Scazzocchio, C.
Cell 27, 4-11, 1981
A:Title: The mosaic organization of the apocytochrome b gene of Aspergillus nidulans rev
A:Reference number: AA0157; MUID:82115341; PMID:7034966
A:Accession: AA0157
A:Molecule type: DNA
A:Residues: 1-387 <WAR>
A:Cross-references: UNIPROT:P00161; UNIPARC:UPI00001289B9; GB:J01389; GB:V00651; GB:V006
A:Experimental source: imperfect stage
C:Genetics:
A:Gene: cobA
A:Genome: mitochondrion
A:Genetic code: SGC3
A:Introns: 169/2

C;Function:

A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released

A;Pathway: oxidative phosphorylation; respiratory chain

C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion F;10-341/Domain: cytochrome b homology <CBH>

F;10-211/Domain: cytochrome b6 homology <CB6>

F;35-51/Domain: transmembrane #status predicted <TM1>

F;80-98/Domain: transmembrane #status predicted <TM2>

F;118-134/Domain: transmembrane #status predicted <TM3>

F;179-201/Domain: transmembrane #status predicted <TM4>

F;223-341/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>

F;231-247/Domain: transmembrane #status predicted <TM5>

F;290-306/Domain: transmembrane #status predicted <TM6>

F;325-345/Domain: transmembrane #status predicted <TM7>

F;355-371/Domain: transmembrane #status predicted <TM8>

F;82_183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted F;96_197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted.

Query Match 5.1%; Score 6; DB 1; Length 387;
Best Local Similarity 100.0%; Pred.No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 189 VLAALA 194

RESULT 605

T35425

A;Molecule type: DNA

A;Residues: 1-387 <OI>

A;Cross-references: UNIPROT:Q9X7Y0; UNIPARC:UPI00000DAF5A; EMBL:AL049485; PIDN:CAB39718.

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCOEBS:SC6A5.34

C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv3272

Query Match 5.1%; Score 6; DB 2; Length 387;
Best Local Similarity 100.0%; Pred.No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
| | | | |
Db 182 GVLAAL 187

RESULT 606

S72995

A;Molecule type: DNA

A;Residues: 1-388 <SMI>

A;Cross-references: UNIPROT:P38056; UNIPARC:UPI00001258F7; EMBL:U00020; NID:g467102; PII C;Genetics:

alanine racemase (EC 5.1.1.1) - Mycobacterium leprae

N;Alternate names: B229 C3_243 protein

C;Species: Mycobacterium leprae

C;Date: 19-Mar-1997 #sequence_revision 23-May-1997 #text_change 09-Jul-2004

C;Accession: S72995

R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993

A;Description: Mycobacterium leprae cosmid B229.

A;Reference number: S72588

A;Accession: S72995

A;Molecule type: DNA

A;Residues: 1-388 <SMI>

A;Cross-references: UNIPROT:P38056; UNIPARC:UPI00001258F7; EMBL:U00020; NID:g467102; PII C;Genetics:

A:Gene: alr
A:Start codon: TTG
C:Superfamily: alanine racemase
C:Keywords: isomerase; phosphoprotein; pyridoxal phosphate
F:44/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 57 LAALAA 62

RESULT 607
AH1950
carbamoyl phosphate synthase small chain [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH1950
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasa moto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH1950
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-388 <KUR>
A:Cross-references: UNIPROT:Q8YXQ7; UNIPARC:UPI0000126F30; GB:BA0000019; PIDN:BA073112.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr1155
C:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) small chain; carbamoyl

Query Match 5.1%; Score 6; DB 2; Length 388;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EYVTS 13
|||||
DB 164 EYVTS 169

RESULT 608
S62597
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Mycena viridimarginata mitochondrion
C:Species: mitochondrion Mycena viridimarginata
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C:Accession: S62597
R:Kraicz, P.; Haase, U.; Gencic, S.; Flindt, S.; Anke, T.; Brandt, U.; von Jagow, G.
Eur. J. Biochem. 235, 54-63, 1996
A:Title: The molecular basis for the natural resistance of the cytochrome bc(1) complex to fungicides
A:Reference number: S62595; MUID:96202917; PMID:8631367
A:Accession: S62597
A>Status: nucleic acid sequence not shown
A:Molecule type: nucleic acid
A:Residues: 1-389 <KRA>
A:Cross-references: UNIPROT:Q36445; UNIPARC:UPI000008DAED; EMBL:X87998; NID:g887554; PIDN:BA073112.1;
C:Genetics:
A:Gene: cyt b
A:Genome: mitochondrion
A:Genetic code: SGC3
A:Introns: 122/3; 131/3; 163/3; 200/3; 289/2
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F:10-341/Domain: cytochrome b homology <CBH>
F:10-210/Domain: cytochrome b6 homology <CB6>
F:35-51/Domain: transmembrane #status predicted <TM1>
F:80-98/Domain: transmembrane #status predicted <TM2>
F:118-134/Domain: transmembrane #status predicted <TM3>

F:179-201/Domain: transmembrane #status predicted <TM4>
F:222-341/Domain: plastoquinol-plastocyanin reductase 17K protein homology <17K>
F:230-246/Domain: transmembrane #status predicted <TM5>
F:289-305/Domain: transmembrane #status predicted <TM6>
F:325-345/Domain: transmembrane #status predicted <TM7>
F:355-371/Domain: transmembrane #status predicted <TM8>
F:82-183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F:96-197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAAAL 25
|||||
DB 189 VLAAAL 194

RESULT 609
G75341
alanine dehydrogenase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: G75341
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodeon, R.J.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75341
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <WHI>
A:Cross-references: UNIPROT:Q9RT70; UNIPARC:UPI00000D3EF5; GB:AE002028; GB:AE000513; NID:BA073112.1;
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1895
A:Map position: 1
C:Superfamily: alanine dehydrogenase; alanine dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGSV 20
|||||
DB 178 VLLGSV 183

RESULT 610
AH0260
conserved hypothetical phage protein YPO2137 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH0260
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0260
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <KUR>
A:Cross-references: UNIPROT:Q8ZEM5; UNIPARC:UPI00000DCDE2; GB:AL590842; PIDN:CAC90948.1;
C:Genetics:
A:Gene: YPO2137

Query Match 5.1%; Score 6; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 70 LAALAA 75

RESULT 611
S25183
Chloramphenicol resistance protein - Rhodococcus fascians
C:Species: Rhodococcus fascians
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S25183; S21395
R:Desomer, J.; Verescke, D.; Crespi, M.; van Montagu, M.
Mol. Microbiol. 6, 2377-2385, 1992
A:Title: The plasmid-encoded chloramphenicol-resistance protein of Rhodococcus fascians
A:Reference number: S25183; MUID:93023865; PMID:1406276
A:Accession: S25183
A:Molecule type: DNA
A:Residues: 1-391 <DES>
A:Cross-references: UNIPROT:Q52751; UNIPARC:UPI00000892A3; EMBL:Z12001; NID:946157; PIDN
C:Genetics:
A:Gene: cmr
A:Start codon: GTG
C:Superfamily: Streptomyces lividans chloramphenicol resistance protein
C:Keywords: antibiotic resistance; membrane protein

Query Match 5.1%; Score 6; DB 1; Length 391;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 131 VLLGGV 136

RESULT 612
F83233
Probable MFS transporter PA3303 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F83233
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,
L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F83233
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <STO>
A:Cross-references: UNIPROT:Q9HYU1; UNIPARC:UPI00000C59A0; GB:AE004752; GB:AE004091; NID
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: PA3303
C:Superfamily: Streptomyces lividans chloramphenicol resistance protein

Query Match 5.1%; Score 6; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
Db 364 VLAALA 369

RESULT 613
E44490
retrovirus-related reverse transcriptase homolog (clone NVD) - pteromalid wasp (Nasonia
C:Species: Nasonia vitripennis
C>Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: E44490
R:Burke, W.D.; Eickbush, D.G.; Xiong, Y.; Jakubczak, J.; Eickbush, T.H.

Mol. Biol. Evol. 10, 163-185, 1993
A:Title: Sequence relationship of retrotransposable elements R1 and R2 within and betwe
A:Reference number: A44490; MUID:93196484; PMID:8383793
A:Contents: retrotransposable element R1
A:Accession: E44490
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-391 <BUR>
A:Cross-references: UNIPROT:Q03272; UNIPARC:UPI0000179238
A:Note: sequence extracted from NCBI backbone (NCBIP:127239)
C:Superfamily: silkworm pol protein

Query Match 5.1%; Score 6; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
Db 77 IELGGK 82

RESULT 614
AG0226
Probable exported protein YP01858 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AG0226
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0226
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-392 <KUR>
A:Cross-references: UNIPROT:Q8ZF60; UNIPARC:UPI00000DCB39; GB:AL590842; PIDN:CAC90675.1
C:Genetics:
A:Gene: YP01858

Query Match 5.1%; Score 6; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 160 VLLGGV 165

RESULT 615
H75444
branched-chain amino acid ABC transporter, periplasmic amino acid-binding protein - Dei
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: H75444
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75444
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <WHI>
A:Cross-references: UNIPROT:Q9RVJ0; UNIPARC:UPI00000D3DBA; GB:AE001955; GB:AE000513; NID
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1038
A:Map position: 1
C:Superfamily: LIV-binding protein

A;Cross-references: UNIPROT:Q9YER8; UNIPARC:UPI000005DC01; DBJ:AP000060; NID:G5104188;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0618
C;Superfamily: Aeropyrum pernix hypothetical protein APE0618

Query Match 5.1%; Score 6; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 376 VLAALA 381

RESULT 621
S35473
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - hillstream loach (Crossostomus
C;Species: Crossostoma lacustre
C;Date: 03-Feb-1994 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C;Accession: S35473; S60283
R;Tzeng, C.S.; Hui, C.F.; Shen, S.C.; Huang, P.C.
Nucleic Acids Res. 20, 4853-4858, 1992
A;Title: The complete nucleotide sequence of the Crossostoma lacustre mitochondrial gene
A;Reference number: S35462; MUID:93027205; PMID:1408600
A;Accession: S35473
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-398 <TZEL>
A;Cross-references: UNIPROT:P34197; UNIPARC:UPI0000171CEA; EMBL:M91245
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1992
R;Tzeng, C.S.; Shen, S.C.; Huang, P.C.
Bull. Inst. Zool. Acad. Sin. 29, 11-19, 1990
A;Title: Mitochondrial DNA identity of Crossostoma (Homalopteridae, Pisces) from two riv
A;Reference number: S60271
A;Accession: S60283
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-380 <TZEE>
A;Cross-references: UNIPARC:UPI00012898F; GB:M91245; NID:G1381122; PIDN:AAB96823.1; PID
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGCI

A;Description: the net reaction catalyzed by the ubiquinol-cytochrome-c reductase comple
ith two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions rele
A;Pathway: oxidative phosphorylation; respiratory chain
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocquinol
C;Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;
F;11-339/Domain: cytochrome b homology <CBH>
F;11-209/Domain: cytochrome b6 homology <CB6>
F;36-52/Domain: transmembrane #status predicted <TM1>
F;81-99/Domain: transmembrane #status predicted <TM2>
F;117-133/Domain: transmembrane #status predicted <TM3>
F;178-200/Domain: transmembrane #status predicted <TM4>
F;221-339/Domain: plastocquinol-plastocyanin reductase 17K protein homology <17K>
F;229-245/Domain: transmembrane #status predicted <TM5>
F;288-304/Domain: transmembrane #status predicted <TM6>
F;323-343/Domain: transmembrane #status predicted <TM7>
F;353-369/Domain: transmembrane #status predicted <TM8>
F;83,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted
F;97,196/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 288 LGGVLA 293

RESULT 622

B87732
protein W10C8.5 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: B87732
R;anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see website genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A;Accession: B87732
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <STO>
A;Cross-references: UNIPROT:O45011; UNIPARC:UPI000005C910; GB:chr_I; PIDN:AAB97594.1; P
C;Genetics:
A;Gene: W10C8.5
A;Map position: 1
C;Superfamily: creatine kinase; creatine kinase repeat homology

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 16 GVLAAL 21

RESULT 623
AH2542
hypothetical protein all7627 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120b-
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2542
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2542
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <KUR>
A;Cross-references: UNIPROT:Q8Z885; UNIPARC:UPI00000CCDBD; GB:AP003602; PIDN:BA877270.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all7627
A;Genome: plasmid
C;Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MTH7

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLA 21
|||||
Db 356 LGGVLA 361

RESULT 624
C95943
probable choline uptake ABC transporter permease protein SMB21145 [imported] - Sinorhiz
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: C95943
R;Finan, T.M.; Weidner, S.; Wong, K.; Ruhmester, J.; Chain, P.; Vorholter, F.J.; Herna
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing end
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: C95943

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-399 <KUR>
A;Cross-references: UNIPROT:Q92Y98; UNIPARC:UPI00000CB67D; GB:AL591985; PIDN:CAC49211.1;
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, J.; Hymann, R.W.; Jones, T.
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, J.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMD21145
A;Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 89 LAALAA 94

RESULT 625
T03460
A;Title: leucine/isoleucine/valine-binding protein precursor - Rhodobacter capsulatus
C;Species: Rhodobacter capsulatus
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03460
R;Vlcek, C.; Paces, V.; Maltsev, N.; Paces, J.; Haselkorn, R.; Fonstein, M.
Proc. Natl. Acad. Sci. U.S.A. 94, 9384-9388, 1997
A;Title: Sequence of a 189-kb segment of the chromosome of Rhodobacter capsulatus SB1003
A;Reference number: Z14955; MUID:97404404; PMID:9256491
A;Accession: T03460
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-400 <VLC>
A;Cross-references: UNIPROT:Q68027; UNIPARC:UPI00000BBDLE; EMBL:AF010496; NID:g3128256;
C;Genetics:
A;Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
| | | | |
Db 327 LAALAA 332

RESULT 626
C88571
A;Title: C05B5.3 [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C;Accession: C88571
R;Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A;Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biological processes
A;Reference number: A75000; MUID:99069613; PMID:9851916
A;Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/
A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A;Accession: C88571
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-401 <STO>
A;Cross-references: UNIPROT:P34291; UNIPARC:UPI000013B731; GB:Chr_III; PIDN:CAA83596.1;
C;Genetics:
A;Gene: C05B5.3

A;Map position: 3

Query Match 5.1%; Score 6; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 EVTTST 13
| | | | |
Db 126 EVTTST 131

RESULT 627

T34715
A;Title: probable ornithine decarboxylase (EC 4.1.1.17) SC1C3.23 [similarity] - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C;Accession: T34715
R;Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21554
A;Accession: T34715
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-402 <OLI>
A;Cross-references: UNIPROT:O69865; UNIPARC:UPI00000DAC50; EMBL:AL023702; PIDN:CAA19247.
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC08DB:SC1C3.23
C;Superfamily: ornithine decarboxylase
C;Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; polyamine biosynthesis;
F;53/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 60 VLAALA 65

RESULT 628

T15677
A;Title: hypothetical protein C28C12.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T15677
R;Miller, N.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of *C. elegans* cosmid C28C12.
A;Reference number: Z18387
A;Accession: T15677
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-402 <MIL>
A;Cross-references: UNIPROT:Q18279; UNIPARC:UPI0000076C88; EMBL:U40797; NID:g1065916; PIDN:U40797;
A;Experimental source: strain Bristol N2; clone C28C12
C;Genetics:
A;Gene: CESP:C28C12.5
A;Map position: 4
A;Introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

Query Match 5.1%; Score 6; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
| | | | |
Db 9 VLAALA 14

RESULT 629

B43260

A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: A1006
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-405 <KUR>
A;Cross-references: UNIPROT:Q8ZJP4; UNIPARC:UPI00000CD65E; GB:AL590842; PIDN:CAC88915.1;
C;Genetics:
A;Gene: dfp
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 5.1%; Score 6; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GHIELG 43
Db 76 GHIELG 81

RESULT 634
G69499
adenosylhomocysteinease (EC 3.3.1.1) homolog ahcy-2 - Archaeoglobus fulgidus
N;Alternate names: S-adenosyl-L-homocysteine hydrolase
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G69499
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glock, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69499
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-405 <KLE>
A;Cross-references: UNIPROT:O28279; UNIPARC:UPI0000056AE3; GB:AE000964; GB:AE000782; NID
C;Superfamily: S-adenosyl-L-homocysteine hydrolase, archaeal SAHH type
C;Keywords: thioether hydrolase

Query Match 5.1%; Score 6; DB 2; Length 405;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 NWQKLE 112
Db 356 NWQKLE 361

RESULT 635
E72545
hypothetical protein APE1649 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: E72545
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:9910339; PMID:10382966
A;Accession: E72545
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-406 <KAW>
A;Cross-references: UNIPROT:Q9YBE8; UNIPARC:UPI000005E027; DDBJ:AF000062; NID:95105244;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1649

Query Match 5.1%; Score 6; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 345 GVLAAL 350

RESULT 636
A44374
3-carboxy-cis,cis-muconate cycloisomerase (EC 5.5.1.2) - Pseudomonas putida
N;Alternate names: 3-carboxy-cis,cis-muconate lactonizing enzyme
C;Species: Pseudomonas putida
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 09-Jul-2004
C;Accession: A44374
R;Williams, S.E.; Woolridge, E.M.; Ransom, S.C.; Landro, J.A.; Babbitt, P.C.; Kozarich,
Biochemistry 31, 9768-9776, 1992
A;Title: 3-Carboxy-cis,cis-muconate lactonizing enzyme from Pseudomonas putida is homolo
A;Reference number: A44374; MUID:93003135; PMID:1390752
A;Accession: A44374
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <WIL>
A;Cross-references: UNIPROT:P32427; UNIPARC:UPI00001313BA; GB:L17082; NID:9309875; PIDN:
A;Note: sequence is inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:115904, NCBIP:115905)
C;Superfamily: fumarate hydratase
C;Keywords: amidine-lyase; carbon-nitrogen lyase; intramolecular lyase; isomerase; purin

Query Match 5.1%; Score 6; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 QAQVIA 79
Db 340 QAQVIA 345

RESULT 637
AC0971
conserved hypothetical protein STY4064 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AC0971
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AC0971
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-407 <PAR>
A;Cross-references: UNIPARC:UPI0000005A6ED; GB:AL513382; PIDN:CAD03263.1; PID:G16504884;
C;Genetics:
A;Gene: STY4064
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 5.1%; Score 6; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 GHIELG 43
Db 75 GHIELG 80

RESULT 638
C83589
conserved hypothetical protein PA0446 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: C83589
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <STO>
A:Cross-references: UNIPROT:Q9I672; UNIPARC:UPI00000C505C; GB:AE004482; GB:AE004091; NID:20437337
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0446
C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV3272

Query Match 5.1%; Score 6; DB 2; Length 407;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 191 VLAALA 196

RESULT 639
A70594
probable manA protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
A:Accession: A70594
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70594
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-408 <COL>
A:Cross-references: UNIPROT:Q05898; UNIPARC:UPI00000D100E; GB:Z95121; GB:AL123456; NID:98295987
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: manA
C:Superfamily: mannose-6-phosphate isomerase

Query Match 5.1%; Score 6; DB 2; Length 408;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
Db 238 GVLAAL 243

RESULT 640
E75290
probable multidrug-efflux transporter - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
A:Accession: E75290
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M. S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioreistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: E75290
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-410 <WHI>
A:Cross-references: UNIPROT:Q9RG20; UNIPARC:UPI00000D3P8B; GB:AE002062; GB:AE000513; NID:98295987
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2307
A:Map position: 1
C:Superfamily: Streptomyces lividans chloramphenicol resistance protein

Query Match 5.1%; Score 6; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||||
Db 169 LLGGVL 174

RESULT 641
H90312
succinyl-diaminopimelate desuccinylase (dapE) [imported] - *Sulfolobus solfataricus*
C:Species: *Sulfolobus solfataricus*
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 05-Oct-2004
A:Accession: H90312
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.; arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: *Sulfolobus solfataricus* complete genome.
A:Reference number: A99139
A:Accession: H90312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-410 <KUR>
A:Cross-references: UNIPROT:Q97V12; UNIPARC:UPI00000644B7; GB:AE006641; NID:913814771; 98295987
C:Genetics:
A:Gene: dapE
C:Superfamily: Succinyl-diaminopimelate desuccinylase

Query Match 5.1%; Score 6; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
|||||
Db 396 KVLGLL 401

RESULT 642
H70908
hypothetical protein Rv0597c - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
A:Accession: H70908
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70908
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-411 <COL>
A:Cross-references: UNIPROT:Q07781; UNIPARC:UPI00000D0PBD; GB:Z97182; GB:AL123456; NID:98295987
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0597c
C:Superfamily: *Mycobacterium* hypothetical protein Rv2008c

Query Match 5.1%; Score 6; DB 2; Length 411;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;

Matches	6;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	20	VLAALA 25							
DB	12	VLAALA 17							
RESULT 643									
C87686									
ferredoxin reductase [imported] - Caulobacter crescentus									
C;Species: Caulobacter crescentus									
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004									
C;Accession: C87686									
R;Nierman, M.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, K.									
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolof									
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.									
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001									
A;Title: Complete Genome Sequence of Caulobacter crescentus.									
A;Reference number: AB7249; MUID:21173698; PMID:11259647									
A;Accession: C87686									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-412 <STO>									
A;Cross-references: UNIPROT:Q9A2N2; UNIPARC:UPI00000C7AEB; GB:AB005673; NID:gl3425257; E									
C;Genetics:									
A;Gene: CG3525									
C;Superfamily: toluene dioxygenase ferredoxin reductase component									
Query Match 5.1%; Score 6; DB 2; Length 412;									
Best Local Similarity 100.0%; Pred. No. 5.9e+02;									
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	33	CVVIVG 38							
DB	10	CVVIVG 15							
RESULT 644									
AC1045									
probable permease yjeh [imported] - Salmonella enterica subsp. enterica serovar Typhi (s									
C;Species: Salmonella enterica subsp. enterica serovar Typhi									
A;Note: This species has also been called Salmonella typhi									
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 05-Oct-2004									
C;Accession: AC1045									
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher									
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,									
, S.; Moulle, S.; O'Gaora, P.									
Nature 413, 848-852, 2001									
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.									
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar									
A;Reference number: AB0502; MUID:21534947; PMID:11677608									
A;Accession: AC1045									
A;Status: preliminary									
A;Molecule type: DNA									
A;Residues: 1-413 <PAR>									
A;Cross-references: UNIPARC:UPI000005A90P; GB:AL513382; PIDN:CAD06808.1; PID:gl6505458;									
C;Genetics:									
A;Gene: yjeh									
C;Superfamily: ecotropic retrovirus receptor protein									
Query Match 5.1%; Score 6; DB 2; Length 413;									
Best Local Similarity 100.0%; Pred. No. 5.9e+02;									
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	21	LAALAA 26							
DB	32	LAALAA 37							
RESULT 645									
AB2726									
conserved hypothetical protein XF1076 [imported] - Xylella fastidiosa (strain 9a5c)									
C;Species: Xylella fastidiosa									

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C;Accession: A82726

R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing

Nature 406, 151-157, 2000

A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A;Reference number: A82515; MUID:20365717; PMID:10910347

A;Note: for a complete list of authors see reference number A59328 below

A;Accession: A82726

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-413 <SIM>

A;Cross-references: UNIPROT:Q9PER2; UNIPARC:UPI000012E7D2; GB:AE003944; GB:AE003849; NID

A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A. Britton, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigret, E.E.; Martins, E.E.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.C.; Miyaki, C.Y.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.A.; Sawasek Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasek A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A;Reference number: A59328

A;Contents: annotation

A;Genetics:

C;Superfamily: hypothetical protein HI1555

Query Match 5.1%; Score 6; DB 2; Length 413;

Best Local Similarity 100.0%; Pred. No. 5.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 AADYIE 73

|||||

Db 94 AADYIE 99

RESULT 646

AG2407

site-specific DNA-methyltransferase [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 31-Dec-2004

C;Accession: AG2407

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AG2407

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-414 <KUR>

A;Cross-references: UNIPROT:Q9VMV9; UNIPARC:UPI00000CEC6D; GB:BA000019; PIDN:BA076514.1

C;Genetics:

A;Gene: alr4815

C;Superfamily: modification methylase (cytosine-specific), M.ECORII type

Query Match 5.1%; Score 6; DB 2; Length 414;

Best Local Similarity 100.0%; Pred. No. 5.9e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88

|||||

Db 180 KGVVLG 185

RESULT 647

D87448

peptidase, M20/M25/M40 family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 16-Aug-2004
C:Accession: D87448
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87448
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <STO>
A:Cross-references: UNIPROT:Q9A7M4; UNIPARC:UPI00000C743D; GB:AE005673; NID:g13423004; E
C:Gene: CC1605
C:Superfamily: Polate hydrolase G

Query Match 5.1%; Score 6; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 140 LAALAA 145

RESULT 648
F84393
threonine synthase (EC 4.2.3.1) [similarity] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84393
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.
; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, I.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84393
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-415 <STO>
A:Cross-references: UNIPROT:Q9HM07; UNIPARC:UPI0000063B2D; GB:AE004437; NID:g10581834; E
C:Gene: thrCl
C:Superfamily: threonine dehydratase
C:Keywords: carbon-oxygen lyase

Query Match 5.1%; Score 6; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAA 27
|||||
Db 144 AALAA 149

RESULT 649
T37023
probable oxidoreductase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37023
R:Murphy, L.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21619
A:Accession: T37023
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-416 <MUR>

Query Match 5.1%; Score 6; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

A:Cross-references: UNIPROT:Q9RI54; UNIPARC:UPI00000DB3D5; EMBL:AL109989; PIDN:CAB53416
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCJ12.04c

Query Match 5.1%; Score 6; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 124 LAALAA 129

RESULT 650
JC4698
divalent cation resistant determinant protein C - Alcaligenes sp.
C:Species: Alcaligenes sp.
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: JC4698
R:Kunito, T.; Kusano, T.; Oyaizu, H.; Senoo, K.; Kanazawa, S.; Matsumoto, S.
Biosci. Biotechnol. Biochem. 60, 699-704, 1996
A:Title: Cloning and sequence analysis of czc genes in Alcaligenes sp. strain CT14.
A:Reference number: JC4698; MUID:96219090; PMID:8829543
A:Accession: JC4698
A:Molecule type: DNA
A:Residues: 1-417 <KUN>
A:Cross-references: UNIPROT:P94175; UNIPARC:UPI00000AFDCB; DBJ:D67024
C:Comment: This protein is a cation/proton antiporter protein, which determines the res
C:Genetics:
A:Gene: czcc
C:Superfamily: cyae protein

Query Match 5.1%; Score 6; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
|||||
Db 105 IELGGK 110

RESULT 651
AI2852
poly(A) polymerase paps [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 05-Oct-2004
C:Accession: AI2852
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AI2852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <KUR>
A:Cross-references: UNIPROT:Q8UD77; UNIPARC:UPI00001646DD; GB:AE008688; PIDN:AAL43239.1
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: paps
A:Map position: circular chromosome
C:Superfamily: poly(A) polymerase/trRNA nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 418;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: exoF1; Smb20945
 A:Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 421;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 25 LAALAA 30

RESULT 657

AC0346

Probable ABC transporter permease YPO2842 [imported] - *Yersinia pestis* (strain CO92)

C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AC0346
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Sammonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrrell, Nature 413, 523-527, 2001

A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0346

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-421 <KUR>

A:Cross-references: UNIPROT:Q8ZC6; UNIPARC:UPI00000CD987; GB:AL590842; PIDN:CAC92094.1;

C:Genetics:

A:Gene: YPO2842

Query Match 5.1%; Score 6; DB 2; Length 421;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 LQRAQ 95
 |||||
 DB 161 LQRAQ 166

RESULT 658

AI3523

Glycerol-3-phosphate-binding periplasmic protein precursor BMEII0115 [imported] - *Brucella melitensis*

C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 C:Accession: AI3523
 R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, P.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*

A:Reference number: AD3252; PMID:11756688

A:Accession: AI3523

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-421 <KUR>

A:Cross-references: UNIPROT:Q8YDR0; UNIPARC:UPI0000058362; GB:AE008918;

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEII0115

A:Map position: II

Query Match 5.1%; Score 6; DB 2; Length 421;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGK 45
 |||||

DB 171 IELGK 176

RESULT 659

H69323

translation initiation factor aIF-2 gamma chain - *Archaeoglobus fulgidus*

C:Species: *Archaeoglobus fulgidus*

C:Date: 26-Aug-1999 #sequence_revision 26-Aug-1999 #text_change 09-Jul-2004

C:Accession: H69323

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69323

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-424 <KLE>

A:Cross-references: UNIPROT:Q29663; UNIPARC:UPI000005705E; GB:AE001064; GB:AE000782; N

C:Superfamily: translation initiation factor eIF-2 gamma chain; translation elongation

C:Keywords: GTP binding; nucleotide binding; P-loop; protein biosynthesis

F26-166/Domain: translation elongation factor Tu homology <ETU>

F32-39/Region: nucleotide-binding motif A (P-loop)

F163-166/Region: GTP-binding NKXD motif

Query Match 5.1%; Score 6; DB 1; Length 424;

Best Local Similarity 100.0%; Pred. No. 6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||

DB 247 LLGGVL 252

RESULT 660

S34449

transcription factor AP-2 - African clawed frog

C:Species: *Xenopus laevis* (African clawed frog)

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C:Accession: S34449

R:Winning, R.S.; Shea, L.J.; Marcus, S.J.; Sargent, T.D.

Nucleic Acids Res. 19, 3709-3714, 1991

A:Title: Developmental regulation of transcription factor AP-2 during *Xenopus laevis* em

A:Reference number: S34449; MUID:91305120; PMID:1852613

A:Accession: S34449

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-425 <WIN>

A:Cross-references: UNIPROT:Q91867; UNIPARC:UPI00000FB3DA; EMBL:M59455; NID:G214835; PII

C:Superfamily: transcription factor AP-2

Query Match 5.1%; Score 6; DB 2; Length 425;

Best Local Similarity 100.0%; Pred. No. 6.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||

DB 237 LLGGVL 242

RESULT 661

JC5086

polytopic cytoplasmic membrane protein - *Aeromonas salmonicida*

C:Species: *Aeromonas salmonicida*

C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 05-Oct-2004

C:Accession: JC5086

R:Noonan, B.; Trust, T.J.

Gene 175, 127-131, 1996

A:Title: An *Aeromonas salmonicida* gene required for the establishment of infection in ra

A:Reference number: JC5086; MUID:97074661; PMID:8917088

A;Accession: JC5086
A;Molecule type: DNA
A;Residues: 1-426 <NOO>
A;Cross-references: UNIPROT:Q44255; UNIPARC:UPI0000083399; GB:L47259; NID:G984559; PIDN:
C;Comment: This protein is involved in translocation across the cytoplasmic membrane.
C;Genetics:
A;Gene: asoB
C;Superfamily: ecotropic retrovirus receptor protein

Query Match 5.1%; Score 6; DB 2; Length 426;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVVL 21
|||||
DB 221 LGGVVL 226

RESULT 662
A40735
TGF beta homolog dal-1 - chicken
C;Species: Gallus gallus (Chicken)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A40735
R;Baeler, K.; Edlund, T.; Jessell, T.M.; Yamada, T.
Cell 73, 687-702, 1993
A;Title: Control of cell pattern in the neural tube: regulation of cell differentiation
A;Reference number: A40735; MUID:93272310; PMID:7916656
A;Accession: A40735
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-427 <BAS>
A;Cross-references: UNIPROT:P34822; UNIPARC:UPI0000129905; GB:L12032; NID:G304379; PIDN:
A;Experimental source: spinal cord
A;Note: sequence extracted from NCBI backbone (NCBIN:132680, NCBI:P:132681)
C;Superfamily: inhibin

Query Match 5.1%; Score 6; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
DB 5 GVLAAL 10

RESULT 663
A87280
conserved hypothetical protein CC0250 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87280
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87280
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-427 <STO>
A;Cross-references: UNIPROT:Q9ABH7; UNIPARC:UPI000006FA6; GB:AE005673; NID:g13421381; H
C;Genetics:
A;Gene: CC0250
C;Superfamily: Rickettsia prowazekii hypothetical protein RP681

Query Match 5.1%; Score 6; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

Db 25 LAALAA 30

RESULT 664
F71713
glycerol-3-phosphate transporter (glpT) RP054 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: F71713
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alenmark, U.
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: F71713
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-429 <AND>
A;Cross-references: UNIPROT:Q9ZE92; UNIPARC:UPI00000D3765; GB:AJ235270; GB:AJ235269; NID:
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: glpT; RP054
C;Superfamily: hexose phosphate transport protein uhpt

Query Match 5.1%; Score 6; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVVL 21
|||||
DB 413 LGGVVL 418

RESULT 665
AF0681
probable voltage gated chloride channel protein STV1574 [imported] - Salmonella enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AF0681
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Stelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0681
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-429 <PAR>
A;Cross-references: UNIPARC:UPI0000059E7C; GB:AL513382; PIDN:CAD01823.1; PID:g16502667;
C;Genetics:
A;Gene: STV1574
C;Superfamily: hypothetical protein all0855

Query Match 5.1%; Score 6; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 138 LAALAA 143

RESULT 666
A65165
pantothenate metabolism flavoprotein dfp - Escherichia coli (strain K-12)
N;Alternate names: probable aspartate i-decarboxylase activase
C;Species: Escherichia coli
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C;Accession: A65165
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of *Escherichia coli* K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: A65165
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-430 <BLAT>
A;Cross-references: UNIPROT:Q8XDA2; UNIPARC:UPI000000D03A6; GB:AB000441; GB:U000096; NID:9
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: dfp
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 5.1%; Score 6; DB 1; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GHIELG 43
Db 98 GHIELG 103

RESULT 667
A87008
hypothetical protein arca [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
A;Accession: A87008
R;Cole, S.T.; Eiglmier, K.; Parthill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: A87008
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <STO>
A;Cross-references: UNIPROT:Q9CC13; UNIPARC:UPI0000125F44; GB:AL450380; NID:gl3092893; F
C;Genetics:
A;Gene: arca
C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carbox

Query Match 5.1%; Score 6; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 34 LAALAA 39

RESULT 668
A90020
preprotein translocase SecY subunit [imported] - *Staphylococcus aureus* (strain N315)
C;Species: *Staphylococcus aureus*
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
A;Accession: A90020
R;Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: A90020
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <KUR>
A;Cross-references: UNIPROT:Q9S39; UNIPARC:UPI0000054BFA; GB:BA000018; PID:gl3702030; F
A;Experimental source: strain N315
C;Genetics:
A;Gene: secY

C;Superfamily: preprotein translocase secY

Query Match 5.1%; Score 6; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
Db 216 KVLGLL 221

RESULT 669

C86040
pantothenate metabolism flavoprotein Dfp [imported] - *Escherichia coli* (strain O157:H7)
C;Species: *Escherichia coli*
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
A;Accession: C86040
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C86040
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <STO>
A;Cross-references: UNIPROT:Q8XDA2; UNIPARC:UPI00000D03A6; GB:AE005174; NID:gl25183395;
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: dfp
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 5.1%; Score 6; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GHIELG 43
Db 98 GHIELG 103

RESULT 670

B91193
pantothenate metabolism flavoprotein Dfp [imported] - *Escherichia coli* (strain O157:H7)
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
A;Accession: B91193
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B91193
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-430 <HAY>
A;Cross-references: UNIPROT:Q8XDA2; UNIPARC:UPI00000D03A6; GB:BA000007; PIDN:BA037937.1
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: EC54514
C;Superfamily: pantothenate metabolism flavoprotein dfp

Query Match 5.1%; Score 6; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GHIELG 43
Db 98 GHIELG 103

RESULT 671

T34927

probable oxidoreductase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34927
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A:Reference number: Z21562
A:Accession: T34927
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-430 <SES>
A:Cross-references: UNIPROT:O69945; UNIPARC:UPI00000DAC8A; EMBL:AL023862; PIDN:CAA19628.
C:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCORDB:SC3F9.05

Query Match 5.1%; Score 6; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 130 LAALAA 135

RESULT 672
T36482
probable aminopeptidase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36482
R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21608
A:Accession: T36482
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-432 <SAU>
A:Cross-references: UNIPROT:Q9XA76; UNIPARC:UPI0000125C48; EMBL:AL096822; PIDN:CAB46924.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCORDB:SCG3D.02

Query Match 5.1%; Score 6; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVVL 21
|||||
DB 282 LGGVVL 287

RESULT 673
F84215
aminopeptidase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84215
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, I.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84215
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <STO>
A:Cross-references: UNIPROT:Q9HRR8; UNIPARC:UPI0000063685; GB:AE004437; NID:g10580170; F
C:Genetics:
A:Gene: ywad

Query Match 5.1%; Score 6; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 390 LAALAA 395

RESULT 674
A29626
apolipoprotein B - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C:Accession: A29626
R:Kirchgesner, T.G.; Heinzmann, C.; Svenson, K.L.; Gordon, D.A.; Nicotia, M.; Lebherz,
Gene 59, 241-251, 1987
A:Title: Regulation of chicken apolipoprotein B: cloning, tissue distribution, and estr
A:Reference number: A29626; MUID:88137960; PMID:3436530
A:Accession: A29626
A:Molecule type: mRNA
A:Residues: 1-433 <KIR>
A:Cross-references: UNIPROT:P11682; UNIPARC:UPI0000125C16; GB:M18421; NID:g211153; PIDN:
C:Genetics:
A:Gene: apob
C:Keywords: chylomicron; lipid binding; lipoprotein; VLDL

Query Match 5.1%; Score 6; DB 2; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 KVLGLL 90
|||||
DB 203 KVLGLL 208

RESULT 675
AF3481
nicotinate phosphoribosyltransferase (EC 2.4.2.11) [imported] - Brucella melitensis (str
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AF3481
R:DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AF3481
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-434 <KUR>
A:Cross-references: UNIPROT:Q8YEP2; UNIPARC:UPI000005820F; GB:AE008917; PIDN:AAL53017.1,
C:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11836
A:Map position: 1
C:Superfamily: nicotinate phosphoribosyltransferase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 5.1%; Score 6; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 247 VLAALA 252

RESULT 676
G97629
pola polymerase (pcnB) rp015 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 05-Oct-2004

C:Accession: G97629
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* strain A77359; PMID:21608551; PMID:11743194
 A:Reference number: A97359; MUID:21608551; PMID:11743194
 A:Accession: G97629
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-435 <KUR>
 A:Cross-references: UNIPROT:Q8U77; UNIPARC:UPI00000D1E29; GB:AE007869; PIDN:AAK87992.1;
 C:Genetics:
 A:Gene: AGR C 4092
 A:Map position: circular chromosome
 C:Superfamily: poly(A) polymerase/tRNA nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 294 LAALAA 299

RESULT 677
 A:5305
 adenylosuccinate lyase - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: A75305
 R:White, O.; Eisen, J.A.; Heideberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: A75305
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-435 <WHI>
 A:Cross-references: UNIPROT:Q9RSB6; UNIPARC:UPI0000132AC3; GB:AE002051; GB:AE000513; NID
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2178
 A:Map position: 1
 C:Superfamily: fumarate hydratase

Query Match 5.1%; Score 6; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
 |||||
 DB 228 VLAALA 233

RESULT 678
 B82753
 D-amino acid dehydrogenase subunit XF0851 [imported] - *Xylella fastidiosa* (strain 9a5c)
 C:Species: *Xylella fastidiosa*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Aug-2004
 C:Accession: B82753
 R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequencing
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A:Reference number: B82515; MUID:20365717; PMID:10910347
 A>Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82753
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-435 <SIM>
 A:Cross-references: UNIPROT:Q9PF27; UNIPARC:UPI0000128DPF; GB:AE003925; GB:AE003849; NID

A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carter, J.; as-Neto, E.; Docena, C.; El-Dorry, H.; Pacincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Proh J.D.; Junqueira, M.D.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasa A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0851
 C:Superfamily: Sarcosine oxidase

Query Match 5.1%; Score 6; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 207 LAALAA 212

RESULT 679
 AC1410
 cellobiose phosphotransferase enzyme IIC component homolog lmo2684 [imported] - *Listeria monocytogenes*
 C:Species: *Listeria monocytogenes*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AC1410
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland A:Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1410
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-435 <GLA>
 A:Cross-references: UNIPROT:Q8Y3Z4; UNIPARC:UPI0000055701; GB:NC_003210; PIDN:CAD00897.1
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo2684
 C:Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depen

Query Match 5.1%; Score 6; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 6.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLLA 23
 |||||
 DB 393 GGVLLA 398

RESULT 680
 AB1786
 cellobiose phosphotransferase enzyme IIC component homolog lin2832 [imported] - *Listeria monocytogenes*
 C:Species: *Listeria monocytogenes*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C:Accession: AB1786
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A>Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1786
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-435 <GLA>
A/Cross-references: UNIPROT:Q927F7; UNIPARC:UPI00000CC9B7; GB:AL592022; PIDN:CAC98058.1;
A/Experimental source: strain Clp11262
C/Genetics:
A/Gene: lin2832
C/Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
|||||
Db 393 GGVLA 398

RESULT 681
C95975
hypothetical protein SPAC4F10.08 - fission yeast (*Schizosaccharomyces pombe*)
C/Species: *Schizosaccharomyces pombe*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T38812
R/Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A/Reference number: Z31813
A/Accession: T38812
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-436 <CON>
A/Cross-references: UNIPROT:O36020; UNIPARC:UPI000006B48E; EMBL:Z98980; PIDN:CAB11711.1;
A/Experimental source: strain 972h-; cosmid c4F10
C/Genetics:
A/Gene: SPDB:SPAC4F10.08
A/Map position: 1
A/Introns: 324/1
C/Superfamily: *Schizosaccharomyces pombe* hypothetical protein SPAC4F10.08

Query Match 5.1%; Score 6; DB 2; Length 436;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGK 45
|||||
Db 387 IELGK 392

RESULT 684
JC5021
platelet-activating factor-acetylhydrolase (EC 3.1.1.-) precursor - guinea pig
N/Alternate names: PAF-acetylhydrolase
C/Species: *Cavia porcellus* (guinea pig)
C/Date: 30-Sep-1993 #sequence_revision 21-Jan-1997 #text_change 05-Oct-2004
C/Accession: JC5021; PC4207
R/Karasawa, K.; Kuge, O.; Kawasaki, K.; Nishijima, M.; Nakano, Y.; Tomita, M.; Yokoyama, J. Biochem. 120, 838-844, 1996
A/Title: Cloning, expression and characterization of plasma platelet-activating factor-a
A/Reference number: JC5021; MUID:97103479; PMID:8947850
A/Accession: JC5021
A/Molecule type: DNA
A/Residues: 1-436 <KAR1>
A/Cross-references: UNIPROT:P70683; UNIPARC:UPI00001311ED; DBJ:D67037; NID:gl644228; PI
A/Accession: PC4207
A/Molecule type: protein
A/Residues: 123-129; 134-139; 208-217; 258-264; 332-337; 341-345; 346-361; 373-384; 385-392 <KAR
A/Cross-references: UNIPARC:UPI00001799F8; UNIPARC:UPI00001799F9; UNIPARC:UPI00001799FA,
9FF; UNIPARC:UPI0000179A00
A/Experimental source: liver
C/Comment: This enzyme converts platelet-activating factor to an inactive metabolite lys
C/Superfamily: platelet-activating factor acetylhydrolase precursor
C/Keywords: glycoprotein; hydrolase

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29
|||||
Db 190 LAAYCL 195

RESULT 682
AD3236
hypothetical protein rbsB [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) pl
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A/Accession: AD3236
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

A>Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1786
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-435 <GLA>
A/Cross-references: UNIPROT:Q927F7; UNIPARC:UPI00000CC9B7; GB:AL592022; PIDN:CAC98058.1;
A/Experimental source: strain Clp11262
C/Genetics:
A/Gene: lin2832
C/Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
|||||
Db 393 GGVLA 398

RESULT 681
C95975
probable polysaccharide polymerase, similar to Wzy protein [imported] - *Sinorhizobium me*
C/Species: *Sinorhizobium meliloti*
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: C95975
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A/Reference number: A95842; MUID:21396508; PMID:11481431
A/Accession: C95975
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-435 <KUR>
A/Cross-references: UNIPROT:Q02729; UNIPARC:UPI000012A37E; GB:AL591985; PIDN:CAC49467.1;
A/Experimental source: strain 1021, megaplasmid pSymB
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pla, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welle, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: exoQ; SMD20944
A/Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29
|||||
Db 190 LAAYCL 195

RESULT 682
AD3236
hypothetical protein rbsB [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) pl
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A/Accession: AD3236
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29
|||||
Db 190 LAAYCL 195

RESULT 682
AD3236
hypothetical protein rbsB [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) pl
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A/Accession: AD3236
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

A>Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1786
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-435 <GLA>
A/Cross-references: UNIPROT:Q927F7; UNIPARC:UPI00000CC9B7; GB:AL592022; PIDN:CAC98058.1;
A/Experimental source: strain Clp11262
C/Genetics:
A/Gene: lin2832
C/Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLA 23
|||||
Db 393 GGVLA 398

RESULT 681
C95975
probable polysaccharide polymerase, similar to Wzy protein [imported] - *Sinorhizobium me*
C/Species: *Sinorhizobium meliloti*
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: C95975
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A/Reference number: A95842; MUID:21396508; PMID:11481431
A/Accession: C95975
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-435 <KUR>
A/Cross-references: UNIPROT:Q02729; UNIPARC:UPI000012A37E; GB:AL591985; PIDN:CAC49467.1;
A/Experimental source: strain 1021, megaplasmid pSymB
R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pla, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.P.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Welle, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: exoQ; SMD20944
A/Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29
|||||
Db 190 LAAYCL 195

RESULT 682
AD3236
hypothetical protein rbsB [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) pl
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A/Accession: AD3236
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

Query Match 5.1%; Score 6; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29
|||||
Db 190 LAAYCL 195

RESULT 682
AD3236
hypothetical protein rbsB [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) pl
C/Species: *Agrobacterium tumefaciens*
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A/Accession: AD3236
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193

F1:1-21/Domain: signal sequence #status predicted <SIG>
F12:436/Product: platelet-activating factor-acetylhydrolase #status predicted <MAT>
F176,200,324/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 436;
Best Local Similarity 100.0%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 EBCSQA 68
DB 225 EBCSQA 230

RESULT 685
A31752
transcription factor AP-2A - human
N:Alternate names: enhancer-binding protein AP-2
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A31752
R:Williams, T.; Admon, A.; Luescher, B.; Tjian, R.
Genes Dev. 2, 1557-1569, 1988
A:Title: Cloning and expression of AP-2, a cell-type-specific transcription factor that
A:Reference number: A31752; MUID:89107991; PMID:3063603
A:Accession: A31752
A:Molecule type: mRNA
A:Residues: 1-437 <WIL>
A:Cross-references: UNIPROT:P05549; UNIPARC:UPI0000125BC5; GB:X52611; GB:Y00229; NID:928
C:Genetics:
A:Gene: GDB:TFAP2A; AP-2; AP2TF; TFAP2
A:Cross-references: GDB:128106; OMIM:107580
A:Map position: 6pter-6p22.3
C:Superfamily: transcription factor AP-2
C:Keywords: alternative splicing; DNA binding; homodimer; transcription factor

Query Match 5.1%; Score 6; DB 1; Length 437;
Best Local Similarity 100.0%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 248 LLGGVL 253

RESULT 686
S42111
transcription factor AP-2 - mouse
N:Alternate names: transcription activator AP-2
C:Species: Mus musculus (house mouse)
C:Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 20-Sep-1999
C:Accession: S42111; S37539
R:Moser, M.; Pscherer, A.; Bauer, R.; Imhof, A.; Seegers, S.; Kerscher, M.; Buettner, R.
Nucleic Acids Res. 21, 4844, 1993
A:Title: The complete murine cDNA sequence of the transcription factor AP-2.
A:Reference number: S42111; MUID:94051614; PMID:8233835
A:Accession: S42111
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-437 <MOS>
A:Cross-references: UNIPARC:UPI000027734; EMBL:X74216; NID:9405776; PIDN:CAAS22292.1; PI
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1993
C:Superfamily: transcription factor AP-2

Query Match 5.1%; Score 6; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 248 LLGGVL 253

RESULT 687

H83556
probable MFS transporter PA0703 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83556
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrner, P.; Hickey, M.J.; B.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83556
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-438 <STO>
A:Cross-references: UNIPROT:O915M3; UNIPARC:UPI00000C510E; GB:AE004506; GB:AE004091; N1
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0703
C:Superfamily: citrate utilization determinant

Query Match 5.1%; Score 6; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 292 LLGGVL 297

RESULT 688
B64915
probable membrane protein b1592 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B64915
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-438 <BLAT>
A:Cross-references: UNIPROT:P76175; UNIPARC:UPI000012A161; GB:AE000255; GB:U00096; NID:
A:Experimental source: strain K-12, substrain MG1655
C:Superfamily: hypothetical protein s110855
C:Keywords: transmembrane protein

F:74-90/Domain: transmembrane #status predicted <TM1>
F:121-137/Domain: transmembrane #status predicted <TM3>
F:142-158/Domain: transmembrane #status predicted <TM4>
F:200-216/Domain: transmembrane #status predicted <TM5>
F:243-259/Domain: transmembrane #status predicted <TM6>
F:281-297/Domain: transmembrane #status predicted <TM7>
F:314-330/Domain: transmembrane #status predicted <TM8>
F:339-355/Domain: transmembrane #status predicted <TM9>
F:370-386/Domain: transmembrane #status predicted <TM10>
F:405-421/Domain: transmembrane #status predicted <TM11>

Query Match 5.1%; Score 6; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.2e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 146 LAALAA 151

RESULT 689
G85764
probable chloride channel Z2583 [imported] - Escherichia coli (strain O157:H7, substrain

C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: G85764
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: G85764
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-438 <STO>
A;Cross-references: UNIPROT:O8X794; UNIPARC:UPI000012A160; GB:AE005174; NID:gl12515570; E
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
C;Superfamily: hypothetical protein sll10855

Query Match 5.1%; Score 6; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 146 LAALAA 151

RESULT 690
B90916
probable chloride channel [imported] - Escherichia coli (strain O157:H7, substrain RIMD
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B90916
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinsagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90916
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-438 <HAV>
A;Cross-references: UNIPROT:O8X794; UNIPARC:UPI000012A160; GB:BA000007; PIDN:BA835721.1;
A;Experimental source: strain O157:H7, substrain RIMD 050952
C;Genetics:
A;Gene: ECs2298
C;Superfamily: hypothetical protein sll10855

Query Match 5.1%; Score 6; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 146 LAALAA 151

RESULT 691
G87290
major facilitator family transporter CC0336 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: G87290
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87290
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-438 <STO>

A;Cross-references: UNIPROT:Q9AB95; UNIPARC:UPI000000C6FED; GB:AE005673; NID:gl13421483; P
C;Genetics:
A;Gene: CC0336

Query Match 5.1%; Score 6; DB 2; Length 438;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|||||
Db 294 LGGVLA 299

RESULT 692
T31071
conserved hypothetical protein - Anabaena sp. (fragment)
C;Species: Anabaena sp.
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2003
C;Accession: T31071
R;Khudyakov, I.; Wolk, C.P.
J. Bacteriol. 179, 6971-6978, 1997
A;Title: HetC, a gene coding for a protein similar to bacterial ABC protein exporters, i
A;Reference number: Z20967; MUID:98037493; PMID:9371442
A;Accession: T31071
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-439 <KHU>
A;Cross-references: UNIPARC:UPI00000B4938; EMBL:U55386; NID:g2828797; PID:g2828799; PIDN
A;Experimental source: POC7120 (also ATCC 27893)
C;Superfamily: beta-glucosidase, GBA2 type

Query Match 5.1%; Score 6; DB 2; Length 439;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 236 LAALAA 241

RESULT 693
H69989
lipoprotein homolog ytcQ - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69989
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fritze, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinola,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadai, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: H69989
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-440 <KUN>
A;Cross-references: UNIPROT:O34936; UNIPARC:UPI000006091E; GB:Z99119; GB:AL009126; NID:
A;Experimental source: strain 168
C;Genetics:
A;Gene: ytcQ
C;Superfamily: Bacillus subtilis lipoprotein lplA

Query Match 5.1%; Score 6; DB 2; Length 440;

Best Local Similarity 100.0%; Pred. No. 6.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 17 LGGVLA 22
|||||
DB 7 LGGVLA 12

RESULT 694
E64667
C:Species: Helicobacter pylori (strain 26695)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: E64667
R:Tomb, J.F.; White, O.; Kerlavags, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne-son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:9739467; PMID:9252185
A:Accession: E64667
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-443 <TOM>
A:Cross-references: UNIPROT:Q25793; UNIPARC:UPI00000D3149; GB:AE000624; GB:AE000511; NID
C:Superfamily: Escherichia coli probable transport protein yajr

Query Match 5.1%; Score 6; DB 1; Length 443;
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 83 KGKVLG 88
|||||
DB 328 KGKVLG 333

RESULT 695
F71848
C:Species: Helicobacter pylori (strain J99)
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: F71848
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Marberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori.
A:Reference number: A71800; MUID:99120557; PMID:99233682
A:Accession: F71848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <ARN>
A:Cross-references: UNIPROT:Q92K35; UNIPARC:UPI00000D36P2; GB:AE001538; GB:AE001439; NID
A:Experimental source: strain J99
C:Genetics:
A:Gene: Jhp1107
C:Superfamily: Escherichia coli probable transport protein yajr

Query Match 5.1%; Score 6; DB 2; Length 443;
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 83 KGKVLG 88
|||||
DB 328 KGKVLG 333

RESULT 696
E69130
C:Species: Helicobacterium thermoautotrophicum (strain Delta H)
A:Title: histidine-tRNA synthetase
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: E69130
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-444 <MTH>
A:Cross-references: UNIPARC:UPI0000165B86; GB:AE000811; GB:AE000666; NID:G2621287; PIDN
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH244
C:Superfamily: histidyl-tRNA synthetase; histidine-tRNA ligase homology
C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis
F:26-412/Domain: histidine-tRNA ligase homology <HTL>

Query Match 5.1%; Score 6; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 4 SADLEV 9
|||||
DB 380 SADLEV 385

RESULT 697
D87557
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: D87557
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; Lau, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonitskii, J.; Krumholz, L.; White, O.; Salzberg, S.L.; Shapero, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: D87557
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: UNIPROT:Q9A5G5; UNIPARC:UPI00000C7746; GB:AE005673; NID:gl3424038;
C:Genetics:
A:Gene: CC2485

Query Match 5.1%; Score 6; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 6.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 17 LGGVLA 22
|||||
DB 297 LGGVLA 302

RESULT 698
S67137
C:Species: Saccharomyces cerevisiae
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C:Accession: S67137
R:Boyer, J.; Fairhead, C.; Gaillon, L.; Galisson, F.; Michaux, G.; Thierry, A.; Dujon, P.; submitted to the Protein Sequence Database, July 1996
A:Reference number: S67104
A:Accession: S67137
A:Molecule type: DNA
A:Residues: 1-445 <BOY>
A:Cross-references: UNIPROT:Q08649; UNIPARC:UPI0000052F1A; EMBL:Z75152; NID:gl420555; P
A:Experimental source: strain S288C

C;Genetics:

A;Gene: SGD:ESA1

A;Cross-references: SGD:S0005770; MIPS:YOR244w

A;Map position: 15R

Query Match 5.1%; Score 6; DB 2; Length 445;

Best Local Similarity 100.0%; Pred. No. 6.3e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 DKEVLY 56

|||||

Db 75 DKEVLY 80

RESULT 699

S45112

transcription factor Ap-2 beta - mouse

C;Species: Mus musculus (house mouse)

C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C;Accession: S45112

R;Moser, M.; Buettnner, R.

submitted to the EMBL Data Library, March 1994

A;Description: Isolation of a second AP-2 related transcription factor:AP-2 beta.

A;Reference number: S45112

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-448 <MOS>

A;Cross-references: UNIPROT:O61131; UNIPARC:UPI0000022746; EMBL:X781597; NID:G496638; PID

C;Superfamily: transcription factor AP-2

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 448;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLY 21

|||||

Db 256 LGGVLY 261

RESULT 700

E98303

hypothetical 49.3K protein in idh-deor intergenic region [imported] - Agrobacterium tum

C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004

C;Accession: E98303

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,

A.; Liu, F.; Mollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: E98303

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-450 <KUR>

A;Cross-references: UNIPROT:Q8UAD3; UNIPARC:UPI00000D21E0; GB:AE007870; PIDN:AAK89951.1;

C;Genetics:

A;Gene: AGR_L 2773

A;Map position: linear chromosome

C;Superfamily: nitrilotriacetate monoxygenase

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 450;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

|||||

Db 81 LAALAA 86

RESULT 701

AH2979

nitrilotriacetate monoxygenase [imported] - Agrobacterium tumefaciens (strain C58, Dupd

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AH2979

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AH2979

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-450 <KUR>

A;Cross-references: UNIPROT:Q8UAD3; UNIPARC:UPI00000D21E0; GB:AE008689; PIDN:AAAL44254.1;

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3441

A;Map position: linear chromosome

C;Superfamily: nitrilotriacetate monoxygenase

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 450;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

|||||

Db 81 LAALAA 86

RESULT 702

T40758

hypothetical protein SPBC8D2.13 - fission yeast (Schizosaccharomyces pombe)

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T40758

R;Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.

submitted to the EMBL Data Library, March 1998

A;Reference number: Z21948

A;Accession: T40758

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-451 <LYN>

A;Cross-references: UNIPROT:O43076; UNIPARC:UPI00000698E6; EMBL:AL022072; PIDN:CAAL17828.

A;Experimental source: strain 972h-; cosmid c8D2

C;Genetics:

A;Gene: SPDB:SPBC8D2.13

A;Map position: 2

A;Introns: 25/3; 48/2; 157/3; 403/2

C;Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC8D2.13

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 451;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

|||||

Db 406 VLAALA 411

RESULT 703

T35729

hypothetical protein SC7H1.34c SC7H1.34c - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004

C;Accession: T35729

R;Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, January 1998

A;Reference number: Z21548

A;Accession: T35729

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A:Residues: 1-452 <MUR>
A:Cross-references: UNIPROT:O54193; UNIPARC:UPI00000DAB03; EMBL:AL021411; PIDN:CAA16221.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SC0EDB:SC7H1.34C

Query Match 5.1%; Score 6; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
Db 119 VLAALA 124
|||||

RESULT 704
AG1293
hypothetical RNA methyltransferase homolog lml1751 [imported] - Listeria monocytogenes
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 05-Oct-2004
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1293
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-453 <GLA>
A:Cross-references: UNIPROT:Q8Y6D6; UNIPARC:UPI00000CF1A3; GB:NC_003210; PIDN:CAC959829.1
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lml1751
C:Superfamily: Hypothetical protein HI0333

Query Match 5.1%; Score 6; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQ 58
Db 292 EVLYQQ 297
|||||

RESULT 705
AB1665
hypothetical RNA methyltransferase homolog lin1863 [imported] - Listeria innocua (strain
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 05-Oct-2004
A:Accession: AB1665
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1665
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-453 <GLA>
A:Cross-references: UNIPROT:Q92A07; UNIPARC:UPI00000CC6B6; GB:AL592022; PIDN:CAC97093.1
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1863
C:Superfamily: Hypothetical protein HI0333

Query Match 5.1%; Score 6; DB 2; Length 453;

Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQ 58
Db 292 EVLYQQ 297
|||||

RESULT 706
F75580
probable sugar transporter - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: F75580
R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, H.O.; Venter, J.C.; Fraser, C.M.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75580
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-454 <WHI>
A:Cross-references: UNIPROT:Q9RYN9; UNIPARC:UPI00000D3BF3; GB:AE001863; GB:AE001825; NTI
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0271
A:Map position: 2
C:Superfamily: yaaU protein

Query Match 5.1%; Score 6; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 432 LAALAA 437
|||||

RESULT 707
AF0215
probable 4-hydroxyphenylacetate permease hpaX [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF0215
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0215
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-455 <KUR>
A:Cross-references: UNIPROT:Q8ZFE6; UNIPARC:UPI00000DC845; GB:AL590842; PIDN:CAC90586.1.
C:Genetics:
A:Gene: hpaX

Query Match 5.1%; Score 6; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 206 VLLGGV 211
|||||

RESULT 708
C86860
hypothetical protein glmU [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C;Accession: C86860
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, G.; Artamonov, S.; Zverev, S.; et al. 2001
A;Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* ssp. *lactis* strain 11461
A;Reference number: A86625; MUID:21235186; PMID:11337471
A;Accession: C86860
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-458 <STO>
A;Cross-references: UNIPROT:Q9CEP8; UNIPARC:UPI000000C6B49; GB:AE005176; PID:gl2724917; EMBL:AL035478
A;Experimental source: strain 11461
C;Genetics:
C;Superfamily: N-acetylglucosamine-1-phosphate uridylyltransferase

Query Match 5.1%; Score 6; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 VVIVGH 39
|||||
Db 49 VVIVGH 54
|||||

RESULT 709
C71884
hypothetical protein jhp0817 - *Helicobacter pylori* (strain J99)
C;Species: *Helicobacter pylori*
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
A;Accession: C71884
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; et al. 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric pathogen *Helicobacter pylori*
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: C71884
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-460 <ARN>
A;Cross-references: UNIPROT:Q9ZKW7; UNIPARC:UPI000012FA8D; GB:AE001511; NID:gl3424124; PID:gl3424124
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0817
C;Superfamily: mvin protein

Query Match 5.1%; Score 6; DB 2; Length 460;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
|||||
Db 159 VLLGGV 164
|||||

RESULT 710
E64630
virulence factor mvin protein - *Helicobacter pylori* (strain 26695)
C;Species: *Helicobacter pylori*
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
A;Accession: E64630
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, R.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; et al. 1997
A;Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*
A;Reference number: A64520; MUID:97394467; PMID:9252185
A;Accession: E64630
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-461 <TOM>

C;Cross-references: UNIPROT:Q25551; UNIPARC:UPI000012FA8E; GB:AE000598; GB:AE000511; NID:gl3424124; PID:gl3424124
C;Superfamily: mvin protein

Query Match 5.1%; Score 6; DB 1; Length 461;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
|||||
Db 160 VLLGGV 165
|||||

RESULT 711
D87566
hypothetical protein CC2558 [imported] - *Caulobacter crescentus*
C;Species: *Caulobacter crescentus*
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004
A;Accession: D87566
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.; et al. 2001
A;Title: Complete genome sequence of *Caulobacter crescentus*
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: D87566
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-462 <STO>
A;Cross-references: UNIPROT:Q9A596; UNIPARC:UPI000000C777F; GB:AE005673; NID:gl3424124; PID:gl3424124
C;Genetics:
A;Gene: CC2558
C;Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 5.1%; Score 6; DB 2; Length 462;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 305 LAALAA 310
|||||

RESULT 712
T34841
probable bifunctional synthase /transferase - *Streptomyces coelicolor*
C;Species: *Streptomyces coelicolor*
C;Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
A;Accession: T34841
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.; et al. 1999
A;Title: The complete genome sequence of *Streptomyces coelicolor*
A;Reference number: Z21559
A;Accession: T34841
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-463 <OLI>
A;Cross-references: UNIPROT:Q9Z5B5; UNIPARC:UPI000000DAEC2; EMBL:AL035478; PIDN:CAB36595
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SC2G5.08
C;Superfamily: hypothetical protein b3052

Query Match 5.1%; Score 6; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
|||||
Db 54 AALAA 59
|||||

RESULT 713
H70504
probable GTP-binding protein - *Mycobacterium tuberculosis* (strain H37RV)

C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: H70504
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Cole, S.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70504
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-463 <COL>
A:Cross-references: UNIPROT:O33212; UNIPARC:UPI0000129EFC; GB:Z98268; GB:AL123456; NID:9
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv1713
C:Superfamily: Mycobacterium leprae probable GTP-binding protein; translation elongation
C:Keywords: duplication; GTP binding; nucleotide binding; P-loop
F:33-40/Domain: nucleotide-binding motif A (P-loop)
F:142-145/Region: GTP-binding motif
F:168-170/Region: GTP-binding SAK/L motif #status atypical
F:200-321/Domain: translation elongation factor Tu homology <E2>
F:206-213/Region: nucleotide-binding motif A (P-loop)
F:318-321/Region: GTP-binding NKXD motif
F:351-353/Region: GTP-binding SAK/L motif

Query Match 5.1%; Score 6; DB 2; Length 463;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 181 GVLAAL 186
|||||

RESULT 714
S36501
L2 protein - human papillomavirus type 27
C:Species: human papillomavirus type 27
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S36501
R:Delius, H.; Hofmann, B.
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36501
A:Molecule type: DNA
A:Residues: 1-464
A:Cross-references: UNIPROT:P36755; UNIPARC:UPI00001388PD; EMBL:X74473; NID:g396964; PID
C:Superfamily: papillomavirus L2 protein
C:Keywords: late protein

Query Match 5.1%; Score 6; DB 2; Length 464;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EVTST 13
Db 128 EVTST 133
|||||

RESULT 715
S15626
protein L2 - human papillomavirus type 57
C:Species: human papillomavirus type 57
A:Note: host Homo sapiens (man)
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: S15626
R:Hirsch-Behnam, A.; Delius, H.; de Villiers, E.M.
Virus Res. 18, 81-98, 1990

A:Title: A comparative sequence analysis of two human papillomavirus (HPV) types 2a and
A:Reference number: S15614; MUID:91188699; PMID:1964523
A:Accession: S15626
A:Molecule type: DNA
A:Residues: 1-465 <HR>
A:Cross-references: UNIPROT:P22164; UNIPARC:UPI000013891A; EMBL:X55965; NID:g60882; PID:
C:Superfamily: papillomavirus L2 protein
C:Keywords: late protein

Query Match 5.1%; Score 6; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EVTST 13
Db 129 EVTST 134
|||||

RESULT 716
G82198
RTX toxin transporter VC1447 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: G82198
R:Reidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82198
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <HEI>
A:Cross-references: UNIPROT:Q9X4W5; UNIPARC:UPI00000D4141; GB:AE004223; GB:AE003852; NI
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1447
A:Map position: 1
C:Superfamily: hemolysin secretion protein D; lipoyl/biotin-binding homology

Query Match 5.1%; Score 6; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 64 GVLAAL 69
|||||

RESULT 717
B97213
fAD/FMN-containing dehydrogenase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: B97213
R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl.
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: B97213
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-467 <KUR>
A:Cross-references: UNIPROT:Q97G30; UNIPARC:UPI00000CA56C; GB:AE001437; PIDN:AAK80493.1
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC2542
C:Superfamily: glycolate oxidase chain glcd

Query Match 5.1%; Score 6; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 IELGOK 45
|||||
Db 182 IELGOK 187

RESULT 718

A75145
protein translocase chain (secY) PAB2139 - Pyrococcus abyssi (strain OsaY)
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: A75145
R:Anonymous, Genoscope
A:Submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure
A:Reference number: A75001
A:Accession: A75145
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <KAW>
A:Cross-references: UNIPROT:Q9VIV8; UNIPARC:UPI00001357EE; GB:AJ248284; GB:AL096836; NID:718
A:Experimental source: strain OsaY
C:Genetics:
A:Gene: PAB2139
C:Superfamily: yeast SSH1 protein

Query Match 5.1%; Score 6; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIV 49
|||||
Db 208 GKPAIV 213

RESULT 719

D70646
probable acyl-CoA dehydrogenase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70646
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70646
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-468 <COL>
A:Cross-references: UNIPROT:P95187; UNIPARC:UPI00000D5FDD; GB:Z83867; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: fadB24

Query Match 5.1%; Score 6; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPA 47
|||||
Db 278 LGGKPA 283

RESULT 720

E71184
probable preprotein translocase secY subunit - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C:Accession: E71184

R:Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekimura, M.; Ohnuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, D. Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon Pyrococcus horikoshii
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71184
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-468 <KAW>
A:Cross-references: UNIPROT:O59442; UNIPARC:UPI00001357F0; GB:AP0000007; NID:93236134; PI:100000000
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1754

Query Match 5.1%; Score 6; DB 2; Length 468;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIV 49
|||||
Db 208 GKPAIV 213

RESULT 721

AC0950
Two-component system, response regulator STY3876 [imported] - Salmonella enterica subsp. C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 02-Jun-2003
C:Accession: AC0950
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.; Connor, R.; Davies, R.; Cronin, A.; Davis, P.; Davies, P.; Dowd, L.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; White, N.; Farrar, S.; Connor, R.; Davies, R.; Cronin, A.; Davis, P.; Davies, P.; Dowd, L.; Farrar, S.; Moule, S.; O'Gaora, P.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AC0950
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <PAR>
A:Cross-references: UNIPARC:UPI000005A65E; GB:AL513382; PIDN:CAD03095.1; PID:gl6504732;
C:Genetics:
A:Gene: STY3876
C:Superfamily: response regulator, NtrC type; response regulator homology; RNA polymerase

Query Match 5.1%; Score 6; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 25
|||||
Db 39 VLALAA 44

RESULT 722

SS3024
nitrogen regulation protein I ntrC - Salmonella typhimurium
C:Species: Salmonella typhimurium
C>Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
C:Accession: SS3024
R:Kustu, S.G.
submitted to the EMBL Data Library, March 1995
A:Reference number: SS3022
A:Accession: SS3024
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <KUS>
A:Cross-references: UNIPROT:P41789; UNIPARC:UPI0000170399; EMBL:X85104; NID:g728721; PID:722
A:Experimental source: strain LT2, substrain DB9005
C:Genetics:
A:Gene: ntrC; glng

C:Function:
A:Description: phosphorylated nitrogen regulation protein I (ntrC) activates transcription
A>Note: for transcription activation: assembly of a multimetric ntrC complex at the enhan
C:Superfamily: response regulator, NtrC type; response regulator homology; RNA polymerase
C:Keywords: ATP; DNA binding; P-loop; phosphoprotein; signal transduction; transcription
F:6-115/Domain: response regulator homology <RRH>
F:140-362/Domain: RNA polymerase sigma factor interaction domain homology <SPI>
F:168-175/Region: nucleotide-binding motif A (P-loop) #status atypical
F:235-239/Region: nucleotide-binding motif B
F:54/Binding site: phosphate (asp) (covalent) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||
Db 39 VLAALA 44

RESULT 723
H64906
aldehyde dehydrogenase homolog b1525 - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 15-Mar-2004
C:Accession: H64906
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.: Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64906
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-470 <BLAT>
A:Cross-references: UNIPARC:UPI00001680B3; GB:AE000250; GB:U00096; NID:gl1787801; PIDN:AP
A:Experimental source: strain K-12, substrain MGL655
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
F:38-296/Domain: aldehyde dehydrogenase homology <ALDD>

Query Match 5.1%; Score 6; DB 2; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QQAVIE 101
|||
Db 126 QQAVIE 131

RESULT 724
D90895
probable aldehyde dehydrogenase [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: D90895
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: D90895
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <HAY>
A:Cross-references: UNIPROT:Q8XAZ8; UNIPARC:UPI0000000656; GB:BA000007; PIDN:BA035555.1;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs2132
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QQAVIE 101
|||
Db 126 QQAVIE 131

RESULT 725
D85722
probable aldehyde dehydrogenase yneI [imported] - Escherichia coli (strain O157:H7, sub:
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: D85722
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85722
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-470 <STO>
A:Cross-references: UNIPROT:Q8XAZ8; UNIPARC:UPI0000000656; GB:AE005174; NID:gl2515137;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yneI
C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 470;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QQAVIE 101
|||
Db 126 QQAVIE 131

RESULT 726
H87284
phosphate regulon sensor histidine kinase PhoR [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: H87284
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kol
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87284
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-471 <STO>
A:Cross-references: UNIPROT:Q9ABD9; UNIPARC:UPI0000006FC4; GB:AE005673; NID:gl13421428;
C:Genetics:
A:Gene: CC0289

Query Match 5.1%; Score 6; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||
Db 44 VLAALA 49

RESULT 727
T46217
glutamate-1-semialdehyde aminotransferase - Arabidopsis thaliana
N:Alternate names: protein T8P19.240
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46217

R;Choiane, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, December 1999
A;Reference number: Z23008
A;Accession: T46217
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <SIM>
A;Cross-references: UNIPROT:Q42522; UNIPARC:UPI0000048816; EMBL:AL133315
A;Experimental source: cultivar Columbia; BAC clone T8P19
C;Genetics:
A;Map position: 3
A;Introns: 59/3; 117/3; 332/3
A;Note: T8P19.240
C;Superfamily: ornithine-oxo-acid aminotransferase

Query Match 5.1%; Score 6; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||
Db 118 VLAALA 123

RESULT 728
F82639
Resistance protein mmr homolog XF1765 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 31-Dec-2004
C;Accession: F82639
R;anonymouse, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: F82639
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-472 <SIM>
A;Cross-references: UNIPROT:Q9PCL3; UNIPARC:UPI00000C2816; GB:AE003999; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsuoka, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1765
C;Superfamily: cysteine-tRNA ligase

Query Match 5.1%; Score 6; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||
Db 56 GVLAAL 61

RESULT 730
D64751
amino acid permease ykfD - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
C;Accession: D64751
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D64751
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-475 <BLAT>
A;Cross-references: UNIPARC:UPI0000168016; GB:AE000134; GB:U00096; NID:gl786454; PIDN:AA
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ykfD
C;Superfamily: ecotropic retrovirus receptor protein
C;Keywords: amino acid transport; transmembrane protein; transport protein
F;26-42/Domain: transmembrane #status predicted <TM1>
F;57-73/Domain: transmembrane #status predicted <TM2>
F;133-149/Domain: transmembrane #status predicted <TM3>
F;169-185/Domain: transmembrane #status predicted <TM4>
F;266-272/Domain: transmembrane #status predicted <TM5>
F;295-311/Domain: transmembrane #status predicted <TM6>
F;348-364/Domain: transmembrane #status predicted <TM7>
F;368-384/Domain: transmembrane #status predicted <TM8>
F;418-434/Domain: transmembrane #status predicted <TM9>
F;444-460/Domain: transmembrane #status predicted <TM10>

Query Match 5.1%; Score 6; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 251 LAALAA 256

RESULT 729
C82737
cysteinyI-tRNA synthetase XF0995 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: C82737

R;anonymouse, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: C82737
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-474 <SIM>
A;Cross-references: UNIPROT:Q9PEN3; UNIPARC:UPI000013637D; GB:AE003937; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsuoka, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0995
C;Superfamily: cysteine-tRNA ligase

Query Match 5.1%; Score 6; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||
Db 56 GVLAAL 61

RESULT 730
D64751
amino acid permease ykfD - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 05-Oct-2004
C;Accession: D64751
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: D64751
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-475 <BLAT>
A;Cross-references: UNIPARC:UPI0000168016; GB:AE000134; GB:U00096; NID:gl786454; PIDN:AA
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: ykfD
C;Superfamily: ecotropic retrovirus receptor protein
C;Keywords: amino acid transport; transmembrane protein; transport protein
F;26-42/Domain: transmembrane #status predicted <TM1>
F;57-73/Domain: transmembrane #status predicted <TM2>
F;133-149/Domain: transmembrane #status predicted <TM3>
F;169-185/Domain: transmembrane #status predicted <TM4>
F;266-272/Domain: transmembrane #status predicted <TM5>
F;295-311/Domain: transmembrane #status predicted <TM6>
F;348-364/Domain: transmembrane #status predicted <TM7>
F;368-384/Domain: transmembrane #status predicted <TM8>
F;418-434/Domain: transmembrane #status predicted <TM9>
F;444-460/Domain: transmembrane #status predicted <TM10>

Query Match 5.1%; Score 6; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

RESULT 735
S21455
glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S21455
R:Axelsen, K.B.; Grimm, B.
submitted to the EMBL Data Library, May 1992
A:Reference number: S21454
A:Accession: S21455
A:Molecule type: mRNA
A:Residues: 1-478 <AXE>
A:Cross-references: UNIPROT:P31593; UNIPARC:UPI000016DE7E; EMBL:X65974; NID:g19874; PID:
C:Superfamily: ornithine-oxo-acid aminotransferase
C:Keywords: chloroplast; intramolecular transferase; isomerase

Query Match      5.1%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 VLAALA 25
      |||||
DB      124 VLAALA 129

RESULT 736
S21454
glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S21454
R:Axelsen, K.B.; Grimm, B.
submitted to the EMBL Data Library, May 1992
A:Reference number: S21454
A:Accession: S21454
A:Molecule type: mRNA
A:Residues: 1-478 <AXE>
A:Cross-references: UNIPROT:P31593; UNIPARC:UPI000012BB98; EMBL:X65973; NID:g19872; PID:
C:Superfamily: ornithine-oxo-acid aminotransferase
C:Keywords: chloroplast; intramolecular transferase; isomerase

Query Match      5.1%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 VLAALA 25
      |||||
DB      124 VLAALA 129

RESULT 737
G86646
beta-glucosidase (EC 3.2.1.21) [imported] - Lactococcus lactis subsp. lactis (strain IL1
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
R:Boettgen, G86646
C:Accession: G86646
R:Boitottin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
A:Reference number: A86625; MUID:21235186; PMID:11337471
A:Accession: G86646
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-478 <STO>
A:Cross-references: UNIPROT:Q9CJ31; UNIPARC:UPI00000C6778; GB:AE005176; PID:g12723027; H
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: BglS
C:Superfamily: Agrobacterium beta-glucosidase
C:Keywords: glycosidase; hydrolase

Query Match      5.1%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;

```

```

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      100 IEPIVT 105
      |||||
DB      116 IEPIVT 121

RESULT 738
P90985
mannose-1-phosphate guanylyltransferase [imported] - Escherichia coli (strain O157:H7, su
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: P90985
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyana, K.; Han, C.G.
Gisawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatcori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: P90985
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-478 <HAY>
A:Cross-references: UNIPROT:Q8X7P1; UNIPARC:UPI00000D0A25; GB:BA000007; PIDN:BA036277.1;
C:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: EC2854
C:Superfamily: Helicobacter mannose-6-phosphate isomerase

Query Match      5.1%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 LAALAA 26
      |||||
DB      97 LAALAA 102

RESULT 739
A85831
mannose-1-phosphate guanylyltransferase [imported] - Escherichia coli (strain O157:H7, su
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: A85831
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dialanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85831
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-478 <STO>
A:Cross-references: UNIPROT:Q8X7P1; UNIPARC:UPI00000D0A25; GB:AE005174; NID:g12516243; P
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: cpsB
C:Superfamily: Helicobacter mannose-6-phosphate isomerase

Query Match      5.1%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      21 LAALAA 26
      |||||
DB      97 LAALAA 102

RESULT 740
H64970
mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22) - Escherichia coli (strain
C:Species: Escherichia coli
C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C:Accession: H64970; A55239

```

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of *Escherichia coli* K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: H64970

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-478 <BLAT>

A;Cross-references: UNIPROT:P24174; UNIPARC:UPI0000168104; GB:U0000295; GB:U000096; NID:975102

A;Experimental source: strain K-12, substrain MG1655

R;Aoyama, K.; Hasee, A.M.; Reeves, P.R. Mol. Biol. Evol. 11, 829-838, 1994

A;Title: Evidence for effect of random genetic drift on G+C content after lateral transfer

A;Reference number: A55239; MUID:95115532; PMID:7815923

A;Accession: A55239

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-324, 'L', 326-478 <AOY>

A;Cross-references: UNIPARC:UPI000012EB95; GB:U38473; GB:U11721; NID:G3041811; PIDN:AACT29301

A;Gene: cpsB

C;Superfamily: Helicobacter mannose-6-phosphate isomerase

C;Keywords: nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

DB 97 LAALAA 102

RESULT 741
T29301
hypothetical protein C50P7.10 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29301
R;Johnson, D.; Stellyes, L.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of *C. elegans* cosmid C50P7.
A;Reference number: Z20601
A;Accession: T29301
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-479 <JOH>
A;Cross-references: UNIPROT:Q18758; UNIPARC:UPI0000060P79; EMBL:U41557; PIDN:AAA83309.17
C;Gene: CESP:C50P7.10
A;Introns: 30/1; 118/3; 203/2; 251/3; 299/1; 342/2; 413/1
C;Superfamily: *Agrobacterium* beta-glucosidase

Query Match 5.1%; Score 6; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 IEPIVT 105
|||||

DB 115 IEPIVT 120

RESULT 742
B75102
NADH dehydrogenase (ubiquinone) chain 4 related PAB0806 - *Pyrococcus abyssi* (strain Orsay)

C;Species: *Pyrococcus abyssi*
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B75102
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome structure

A;Reference number: A75001

A;Accession: B75102
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-479 <RAW>
A;Cross-references: UNIPROT:Q9U280; UNIPARC:UPI000003451F; GB:AJ248286; GB:AL096836; NID:975102
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB0806
C;Superfamily: [NiFe]-hydrogenase-3-type complex, large membrane subunit

Query Match 5.1%; Score 6; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
|||||

DB 425 IELGGK 430

RESULT 743
T50726
hypothetical protein 479 [imported] - *Rhodobacter sphaeroides*
C;Species: *Rhodobacter sphaeroides*
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C;Accession: T50726
R;Choudhary, M.; Kaplan, S.
Nucleic Acids Res. 28, 862-867, 2000
A;Title: DNA sequence analysis of the photosynthesis region of *Rhodobacter sphaeroides*
A;Reference number: Z25222; MUID:20115911; PMID:10648776
A;Accession: T50726
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-479 <CHO>
A;Cross-references: UNIPROT:Q9RFD8; UNIPARC:UPI0000082F99; EMBL:AF195122; PIDN:AAF24270
A;Experimental source: strain 2.4.1

Query Match 5.1%; Score 6; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAAAL 25
|||||

DB 138 VLAAAL 143

RESULT 744
F84179
hypothetical protein Vng0189c [imported] - *Halobacterium* sp. NRC-1
C;Species: *Halobacterium* sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84179
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, J.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabluchung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L.
A;Title: Genome sequence of *Halobacterium* species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: F84179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-479 <STO>
A;Cross-references: UNIPROT:Q9HSK2; UNIPARC:UPI000006358A; GB:AE004437; NID:G10579836; PIDN:AAE004437
C;Genetics:
A;Gene: VNG0189C

Query Match 5.1%; Score 6; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||

DB 285 LAALAA 290

RESULT 745

S18447
variant surface glycoprotein ILTat 1.22 - Trypanosoma brucei
C;Species: Trypanosoma brucei
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 09-Jul-2004
C;Accession: S18447
R;Carrington, M.; Miller, N.; Blum, M.; Roditi, I.; Wiley, D.; Turner, M.
J. Mol. Biol. 221, 823-835, 1991
A;Title: Variant specific glycoprotein of Trypanosoma brucei consists of two domains each with a conserved cysteine-rich region
A;Reference number: S18445; MUID:92046037; PMID:1942032
A;Accession: S18447
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-479 <CR>
A;Cross-references: UNIPROT:P26327; UNIPARC:UPI0000138DB5; EMBL:X56765; NID:g10449; PID:S18447
C;Keywords: glycoprotein

Query Match 5.1%; Score 6; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
|||||
Db 93 AALAA 88

RESULT 746

AF0768
mannose-1-phosphate guanylyltransferase (GDP) (EC 2.7.7.22) - Salmonella enterica subsp.
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: This species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF0768
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; et al.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0768
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-480 <PAR>
A;Cross-references: UNIPROT:Q9ETZ0; UNIPARC:UPI0000059C34; GB:AL513382; PIDN:CAD02468.1
A;Gene: manC
C;Superfamily: Helicobacter mannose-6-phosphate isomerase
C;Keywords: nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 99 LAALAA 104

RESULT 747

S16290
mannose-1-phosphate guanylyltransferase - Salmonella typhimurium
C;Species: Salmonella typhimurium
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S16290
R;Stevenson, G.; Lee, S.J.; Romana, L.K.; Reeves, P.R.
Mol. Gen. Genet. 227, 173-180, 1991
A;Title: The cps gene cluster of Salmonella strain LT2 includes a second mannose pathway
A;Reference number: S16289; MUID:91287694; PMID:1712067
A;Accession: S16290
A;Status: preliminary

A;Molecule type: DNA
A;Residues: 1-480 <STE>
A;Cross-references: UNIPROT:P26340; UNIPARC:UPI000012BB99; EMBL:X59886; NID:g47654; PIDN:S18447
C;Superfamily: Helicobacter mannose-6-phosphate isomerase

Query Match 5.1%; Score 6; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 99 LAALAA 104

RESULT 748

H82506
hypothetical protein VCA0066 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: H82506
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; et al.
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: AB2035; MUID:20406833; PMID:10952301
A;Accession: H82506
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-480 <HEI>
A;Cross-references: UNIPROT:Q9KN99; UNIPARC:UPI000000C33P7; GB:AE004349; GB:AE003853; NID:S18447
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0066
A;Map position: 2

Query Match 5.1%; Score 6; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 25
|||||
Db 192 VLALAA 197

RESULT 749

T07034
Glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8) - tomato
N;Alternate names: glutamate 1-semialdehyde aminotransferase
C;Species: Lycopersicon esculentum (tomato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T07034
R;Brander, K.A.; Owtrim, G.W.; Brunold, C.
Plant Physiol. 108, 1748, 1995
A;Title: Isolation of a cDNA (EMBL X85803) encoding a putative chloroplastic isoform of
A;Reference number: Z15026
A;Accession: T07034
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-481 <BRA>
A;Cross-references: UNIPROT:Q40147; UNIPARC:UPI000012BB82; EMBL:L39279; NID:g642910; PID:T07034
A;Experimental source: strain VFNT cherry; young fruit
C;Function:

A;Pathway: porphyrin biosynthesis
C;Superfamily: ornithine-oxo-acid aminotransferase
C;Keywords: intramolecular transferase; isomerase; phosphoprotein; pyridoxal phosphate

Query Match 5.1%; Score 6; DB 2; Length 481;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 25
|||||

Db 127 VLAALA 132

RESULT 750
H69392
4-hydroxybutyrate CoA transferase (cat2-1) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
A:Accession: H69392
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: H69392
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-482 <KLE>
A:Cross-references: UNIPROT:O29120; UNIPARC:UPI0000056E3F; GB:AE001025; GB:AE000782; NID
C:Superfamily: acetyl-CoA hydrolase

Query Match 5.1%; Score 6; DB 2; Length 482;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIELGG 44
|||||
Db 32 HIELGG 37

RESULT 751
AD0041
rhamnulokinase (EC 2.7.1.5) [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
A:Accession: AD0041
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AD0041
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-485 <KUR>
A:Cross-references: UNIPROT:Q8ZJ02; UNIPARC:UPI00000DCA72; GB:AL590842; PIDN:CAC89191.1;
C:Genetics:
A:Gene: rhaB
C:Superfamily: rhamnulokinase
C:Keywords: phosphotransferase

Query Match 5.1%; Score 6; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRAT 94
|||||
Db 297 LLQRAT 302

RESULT 752
T13479
hypothetical protein 34F3.9 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
A:Accession: T13479
R:Valentini, P.; Salles, C.; Campbell, L.; Glover, D.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17685
A:Accession: T13479
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-485 <PHI>
A:Cross-references: UNIPROT:O77271; UNIPARC:UPI0000080EDE; EMBL:AL031583; NID:e1321005;
C:Genetics:
A:Note: EG:34F3.9

Query Match 5.1%; Score 6; DB 2; Length 485;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 10
|||||
Db 127 ADLEVT 132

RESULT 753
S61993
probable membrane protein YOR009w - Yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein O2549; hypothetical protein UNB487
C:Species: Saccharomyces cerevisiae
C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
A:Accession: S61993; S66874; S72142
R:Sterky, F.; Uhlen, M.
submitted to the EMBL Data Library, December 1995
A:Reference number: S61981
A:Accession: S61993
A:Molecule type: DNA
A:Residues: 1-487 <STE>
A:Cross-references: UNIPROT:Q12218; UNIPARC:UPI000006A42C; EMBL:U43491; NID:g1150992; P
R:Petersson, B.; Sterky, F.; Uhlen, M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S6682
A:Accession: S66874
A:Molecule type: DNA
A:Residues: 1-487 <PET>
A:Cross-references: UNIPARC:UPI000006A42C; EMBL:Z74917; NID:g1420103; PID:e251944; PID:
A:Experimental source: strain S288C
R:Sterky, F.; Holmberg, A.; Petersson, B.; Uhlen, M.
Yeast 12, 1091-1095, 1996
A:Title: The sequence of a 30 kb fragment on the left arm of chromosome XV from Sacchar
A:Reference number: S72130; MUID:97051599; PMID:8896276
A:Accession: S72142
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-487 <STW>
A:Cross-references: UNIPARC:UPI000006A42C; EMBL:U43491; NID:g1150992; PIDN:AAC49489.1;
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1995
C:Genetics:
A:Gene: SGD:TIR4
A:Cross-references: SGD:S0005535
A:Map position: 15R
A:Note: YOR009w
C:Keywords: transmembrane protein
F:471-487/Domain: transmembrane #status predicted <TMM>

Query Match 5.1%; Score 6; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 9 LAALAA 14

RESULT 754
F83109
probable phosphate transporter PA4292 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: F83109
R;Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bz
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83109
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-489 <STO>
A;Cross-references: UNIPROT:Q9HWA8; UNIPARC:UPI00000C5C9D; GB:AE004845; GB:AE004091; NID
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4292

Query Match 5.1%; Score 6; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
DB 65 VLLGGV 70

RESULT 755
QARTP2
cytochrome P450 2B1 - rat
N;Alternate names: cytochrome P450 b; cytochrome P450, phenobarbital-inducible
N;Contains: unspecific monooxygenase (EC 1.14.14.1)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 18-Aug-1982 #sequence revision 17-May-1996 #text change 09-Jul-2004
C;Accession: A00176; A54251; A23263; A29298; S03854; A92255; I54796
R;Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982
A;Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phenobar
A;Reference number: A93912; MUID:82222224; PMID:6953431
A;Accession: A00176
A;Molecule type: mRNA
A;Residues: 6-491 <FU>
A;Cross-references: UNIPROT:P00176; UNIPROT:Q64584; UNIPARC:UPI00000619F3; EMBL:J00719;
A;Note: the authors translated the codon GAT for residue 166 as Glu, CTG for residue 292
R;Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982
A;Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobarb
A;Reference number: A93925
A;Contents: annotation
A;Note: the mistranslations shown in reference A93912 are acknowledged
R;Roberts, E.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg, P.
Biochemistry 33, 3766-3771, 1994
A;Title: Identification of active-site peptides from (3)H-labeled 2-ethynylthalene-1
A;Reference number: A54251; MUID:94190899; PMID:8142377
A;Accession: A54251
A;Molecule type: protein
A;Residues: 290-301, 'X' <ROB>
A;Cross-references: UNIPARC:UPI0000171D4D
R;Suwa, Y.; Mizukami, Y.; Sogawa, K.; Fujii-Kuriyama, Y.
J. Biol. Chem. 260, 7980-7984, 1985
A;Title: Gene structure of a major form of phenobarbital-inducible cytochrome P-450 in
A;Reference number: A22363; MUID:85234490; PMID:2989270
A;Accession: A22363
A;Molecule type: DNA
A;Residues: 1-91, 'P', 93-204, 'R', 206-327, 'V', 329-356, 'H', 358-391, 'R', 393-415, 'V', 417-433,
A;Cross-references: UNIPARC:UPI00000E78C9; GB:L00320; NID:G203816; PIDN:AAAA1046.1; PID:
A;Note: the authors translated the codon CAG for residue 57 as Gly, CTT for residue 92 a
as Arg
R;Rangarajan, P.N.; Ravishankar, H.; Padmanaban, G.
Biochem. Biophys. Res. Commun. 144, 258-263, 1987
A;Title: Isolation of a cytochrome P-450e gene variant and characterization of its 5' fl
A;Reference number: A29298; MUID:87213174; PMID:3579906
A;Accession: A29298
A;Status: not compared with conceptual translation

A;Molecule type: DNA
A;Residues: 1-57 <RAN>
A;Cross-references: UNIPARC:UPI00000171D4E
R;Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.
Arch. Biochem. Biophys. 270, 23-32, 1989
A;Title: Antibodies targeted against hypervariable and constant regions of cytochromes I
A;Reference number: S03854; MUID:89192373; PMID:2539047
A;Accession: S03854
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-18;146-160,'E',162-165;166,330-361;362-380;402-423 <OES>
A;Cross-references: UNIPARC:UPI0000171D4F; UNIPARC:UPI0000171D50; UNIPARC:UPI0000171D51;
R;Botelho, L.H.; Ryan, D.E.; Levin, W.
J. Biol. Chem. 254, 5635-5640, 1979
A;Title: Amino acid compositions and partial amino acid sequences of three highly purified
or 3-methylcholanthrene.
A;Reference number: A92255; MUID:79194111; PMID:109438
A;Accession: A92255
A;Molecule type: protein
A;Residues: 1-3, 'F', 5-22 <BOT>
A;Cross-references: UNIPARC:UPI0000171D54
R;Fujii-Kuriyama, Y.; Mizukami, Y.; Taniguchi, T.; Muramatsu, M.
Int. Symp. Princess Takamatsu Cancer Res. Fund 12, 31-40, 1982
A;Title: Molecular cloning and coding nucleotide sequence of complementary DNA of cytoch
A;Reference number: I54796; MUID:83160754; PMID:6300027
A;Accession: I54796
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 6-491 <RES>
A;Cross-references: UNIPARC:UPI00000619F3; GB:M37134; NID:G203784; PIDN:AAC42028.1; PID:
C;Genetics:
A;Gene: CYP2B1
A;Introns: 57/3; 112/1; 162/1; 215/3; 274/3; 322/1; 384/3; 432/1
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo
F;295-458/Domain: Cytochrome P450 homology <P45>
F;302/Active site: Thr #status predicted
F;436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 AVIEPI 103
|||||
DB 102 AVIEPI 107

RESULT 756
QARTP2
cytochrome P450 2B2 - rat
N;Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cyto
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 31-Dec-2004
R;Mizukami, Y.; Sogawa, K.; Suwa, Y.; Muramatsu, M.; Fujii-Kuriyama, Y.
Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983
A;Title: Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver.
A;Reference number: A21162; MUID:83247397; PMID:6306654
A;Accession: A21162
A;Molecule type: DNA
A;Residues: 1-472, 'M', 474-491 <MIZ>
A;Cross-references: UNIPROT:P00176; UNIPROT:Q64579; UNIPARC:UPI000017099;
A;Note: the authors translated the codon AGT for residue 4 as Thr, and ATG for residue 1
R;Frey, A.B.; Waxman, D.J.; Kreibich, G.
J. Biol. Chem. 260, 15253-15265, 1985
A;Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme I
A;Reference number: A00177; MUID:86059379; PMID:3877725
A;Accession: A00177
A;Molecule type: protein
A;Residues: 1-291, 'P', 293-320, 'AE', 323-475, 'D', 477-491 <FRE>
A;Cross-references: UNIPARC:UPI0000171D55

```

F:436/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match          5.1%; Score 6; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

Qy 98 AVIEPI 103
    |||||
Db 102 AVIEPI 107

RESULT 757
B86096
xyllose-proton symport [imported] - Escherichia coli (strain O157:H7;
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Ju
C:Accession: B86096
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.;
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Pot
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B86096
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-491 <STO>
A:Cross-references: UNIPROT:P09098; UNIPARC:UPI000013905F; GB:AE0051
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: xyle
C:Superfamily: glucose transport protein

Query Match          5.1%; Score 6; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

Qy 19 GVLAAL 24
    |||||
Db 452 GVLAAL 457

RESULT 758
F91255
xyllose-proton symport [imported] - Escherichia coli (strain O157:H7;
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Ju
C:Accession: F91255
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yok
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shin
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: F91255
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-491 <HAY>
A:Cross-references: UNIPROT:P09098; UNIPARC:UPI000013905F; GB:BA0000
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECe5014
C:Superfamily: glucose transport protein

Query Match          5.1%; Score 6; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0;

Qy 19 GVLAAL 24
    |||||
Db 452 GVLAAL 457

RESULT 759
A26430

```

xylose transport protein - Escherichia coli (strain K-12)
N/Alternate names: xylose-proton symport protein
C/Species: Escherichia coli
C/Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C/Accession: A26430; A27418; F65210; S00874
R/Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.P.
Nature 325, 641-643, 1987
A/Title: Mammalian and bacterial sugar transport proteins are homologous.
A/Reference number: A93389; MUID:87115869; PMID:3543693
A/Accession: A26430
A/Status: nucleic acid sequence not shown; not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 1-491 <MAL>
A/Cross-references: UNIPROT:P09098; UNIPARC:UPI000013905F
R/Davis, E.O.; Henderson, P.J.P.
J. Biol. Chem. 262, 13928-13932, 1987
A/Title: The cloning and DNA sequence of the gene xylE for xylose-proton symport in Esch
A/Reference number: A27418; MUID:88007632; PMID:2820984
A/Accession: A27418
A/Molecule type: DNA
A/Residues: 1-491 <DAV>
A/Cross-references: UNIPARC:UPI000013905F; GB:J02812; NID:g148282; PIDN:AAA79016.1; PID:
A/Experimental source: strain K12
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of Escherichia coli K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: F65210
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-491 <BLAT>
A/Cross-references: UNIPARC:UPI000013905F; GB:AE000476; GB:U00096; NID:g1790456; PIDN:AA
A/Experimental source: strain K-12, substrain MG1655
R/Francoz, E.; Daese, E.
Nucleic Acids Res. 16, 4097-4109, 1988
A/Title: 3' end of the malEFG operon in E.coli: localization of the transcription termina
A/Reference number: S00776; MUID:88234001; PMID:2836810
A/Accession: S00874
A/Molecule type: DNA
A/Residues: 1-63, 'V', 65-192 <FRA>
A/Cross-references: UNIPARC:UPI000016F2C6; EMBL:X06663; NID:g41953; PIDN:CAA29863.1; PID
C/Genetics:
A/Gene: xyle
A/Map position: 91.4
C/Superfamily: glucose transport protein
C/Keywords: transmembrane protein

Query Match 5.1%; Score 6; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 452 GVLAAL 457

RESULT 760
F96022
conserved hypothetical protein SMB20754 [imported] - Sinorhizobium meliloti (strain 1021
C/Species: Sinorhizobium meliloti
C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C/Accession: F96022
R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo
A/Reference number: A95842; MUID:21396508; PMID:11481431
A/Accession: F96022
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-491 <KUR>
A/Cross-references: UNIPROT:Q92T04; UNIPARC:UPI00000CB886; GB:AL591985; PIDN:CAC49846.1;
A/Experimental source: strain 1021, megaplasmid pSymB

R/Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaunt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A/Reference number: A96039; MUID:21368234; PMID:11474104
A/Contents: annotation
C/Genetics:
A/Gene: SMB20754
A/Genome: plasmid

Query Match 5.1%; Score 6; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAA 27
DB 200 AALAA 205

RESULT 761
AG1384
hypothetical protein lmo2479 [imported] - Listeria monocytogenes (strain EGD-e)
C/Species: Listeria monocytogenes
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AG1384
R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AG1384
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-491 <GLA>
A/Cross-references: UNIPROT:Q8Y4G5; UNIPARC:UPI00000556FD; GB:NC_003210; PIDN:CAD00557.1
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo2479

Query Match 5.1%; Score 6; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 10
DB 283 ADLEVT 288

RESULT 762
A11759
hypothetical protein lin2622 [imported] - Listeria innocua (strain Clip11262)
C/Species: Listeria innocua
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: A11759
R/Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.A.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A/Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,
A/Title: Comparative genomics of Listeria species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: A11759
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-491 <GLA>
A/Cross-references: UNIPROT:Q928B4; UNIPARC:UPI00000CC91F; GB:AL592022; PIDN:CAC97849.1;

A:Experimental source: strain Clipl1262
C:Genetics:
A:Gene: lin2622

Query Match 5.1%; Score 6; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 10
|||||
DB 283 ADLEVT 288

RESULT 763

T06031
hexokinase homolog T28119.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06031
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Ba
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15484
A:Accession: T06031
A:Molecule type: DNA
A:Residues: 1-493 <BEV>
A:Cross-references: UNIPROT:Q9T071; UNIPARC:UPI0000048A89; EMBL:AL035709; GSFDB:GN000062;
A:Experimental source: cultivar Columbia; BAC clone T28119
C:Genetics:
A:Gene: ATSP:T28119.120
A:Map position: 4
A:Introns: 92/2; 142/3; 227/3; 279/3; 312/3; 356/2; 383/3
C:Superfamily: hexokinase; hexokinase homology

Query Match 5.1%; Score 6; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 YIEQAQ 76
|||||
DB 259 YIEQAQ 264

RESULT 764

S76517
hypothetical protein slr0643 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76517
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76517
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-493 <KAN>
A:Cross-references: UNIPROT:Q55722; UNIPARC:UPI00000C0F47; EMBL:D64002; GB:AB001339; NID
A:Note: the nucleotide sequence was submitted to the EMBL data Library, June 1996

Query Match 5.1%; Score 6; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 48 IVPDKK 53
|||||
DB 80 IVPDKK 85

RESULT 765

E75081

alkaline phosphatase (EC 3.1.3.1) IV PAB2366 precursor - Pyrococcus abyssi (strain Osa;
N:Alternate names: apase IV
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 16-Aug-2004
C:Accession: E75081
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome str
A:Reference number: A75001
A:Accession: E75081
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-495 <KAW>
A:Cross-references: UNIPROT:Q9UZV2; UNIPARC:UPI0000034702; GB:AJ248286; GB:AL096836; NTI
A:Experimental source: strain Ozeay
C:Genetics:
A:Gene: phoA; PAB2366
C:Superfamily: Alkaline phosphatase
C:Keywords: phosphoric monoester hydrolase

Query Match 5.1%; Score 6; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
|||||
DB 200 GKVLGL 205

RESULT 766

T14236
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 - Synechococcus sp. (strain PCC 77
C:Species: Synechococcus sp.
A:Variety: strain PCC 7002
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14236
R:Klughammer, B.; Sultemeyer, D.; Badger, M.R.; Price, G.D.
submitted to the EMBL Data Library, April 1997
A:Description: Involvement of ndhF3, ndhD3 and ORF427 genes in high affinity CO2 uptake
A:Reference number: Z17936
A:Accession: T14236
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-498 <KLJ>
A:Cross-references: UNIPROT:Q33750; UNIPARC:UPI00000BEC97; EMBL:U97516; NID:g2232044; P
A:Experimental source: strain PCC 7002
C:Genetics:
A:Note: ndh3
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
DB 237 VLLGGV 242

RESULT 767

S62626
protein disulfide-isomerase (EC 5.3.4.1) - castor bean
C:Species: Ricinus communis (castor bean)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S62626
R:Coughlan, S.J.; Hastings, C.; Winfrey Jr., R.J.
Eur. J. Biochem. 235, 215-224, 1996
A:Title: Molecular characterisation of plant endoplasmic reticulum: identification of p
A:Reference number: S62620; MUID:96202938; PMID:8631332
A:Accession: S62626
A:Status: preliminary
A:Molecule type: mRNA

A;Residues: 1-498 <COU>
A;Cross-references: UNIPROT:Q43116; UNIPARC:UPI00001314C1; EMBL:U41385; NID:g1134967; PI
C;Superfamily: protein disulfide-isomerase; thioredoxin homology
C;Keywords: intramolecular oxidoreductase; isomerase
F;40-128/Domain: thioredoxin homology <TXN>

Query Match 5.1%; Score 6; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 AHQPKG 84
|||||
Db 294 AHQPKG 299

RESULT 768
A53956
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal - human
C;Species: Homo sapiens (man)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: A53956; S21338
R;Mihovilovic, M.; Roses, A.D.
Exp. Neurol. 111, 175-180, 1991
A;Title: Expression of mRNAs in human thymus coding for the alpha3 subunit of a neuronal
A;Reference number: A53956; MUID:91114756; PMID:1989896
A;Accession: A53956
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-503 <MIH>
A;Cross-references: UNIPROT:P32297; UNIPARC:UPI000016ADF3; GB:M37981; NID:g189252; PIDN:
R;Anand, R.; Lindstrom, J.
submitted to the EMBL Data Library, June 1990
A;Description: Nucleotide sequence of the mature human nicotinic acetylcholine receptor
A;Reference number: S21338
A;Accession: S21338
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-503 <ANA>
A;Cross-references: UNIPARC:UPI000016ADAE; EMBL:X53559; NID:g34985; PIDN:CAA37625.1; PID
C;Genetics:
A;Gene: GDB:CHRNA3
A;Cross-references: GDB:125219; OMIM:118503
A;Map position: 15q24-15q24
C;Superfamily: acetylcholine receptor
C;Keywords: neurotransmitter receptor

Query Match 5.1%; Score 6; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 3 LAALAA 8

RESULT 769
S55589
D-nopaline dehydrogenase (EC 1.5.1.19) OoxA - Agrobacterium tumefaciens
C;Species: Agrobacterium tumefaciens
C;Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C;Accession: S55589
R;Zanker, H.; Lurz, G.; Langridge, U.; Langridge, P.; Schroeder, J.
submitted to the EMBL Data Library, February 1994
A;Description: Octopine and nopaline oxidases from Ti plasmids of Agrobacterium tumefaci
A;Reference number: S55578
A;Accession: S55589
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-503 <ZAN>
A;Cross-references: UNIPROT:Q59160; UNIPARC:UPI00000008B3; EMBL:Z30328; NID:g496539; PID
C;Keywords: oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 503;

Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 337 LAALAA 342

RESULT 770
AD3128
choline sulfatase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AD3128
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD3128
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-503 <KUR>
A;Cross-references: UNIPROT:Q8U708; UNIPARC:UPI00000D264F; GB:AE008689; PIDN:AAI45442.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: betC
A;Map position: linear chromosome

Query Match 5.1%; Score 6; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 360 LAALAA 365

RESULT 771
E98159
choline sulfatase (U39940) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: E98159
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, B.;
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: E98159
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-503 <KUR>
A;Cross-references: UNIPROT:Q8U708; UNIPARC:UPI00000D264F; GB:AE007870; PIDN:AAK88799.1;
C;Genetics:
A;Gene: AGR_L_469
A;Map position: linear chromosome

Query Match 5.1%; Score 6; DB 2; Length 503;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 360 LAALAA 365

RESULT 772
T09437
probable aldehyde dehydrogenase (NAD) (EC 1.2.1.3) alda [similarity] - Vibrio cholerae

C:Species: *Vibrio cholerae*
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C:Accession: T09437
 R:Karacolis, D.K.R.; Johnson, J.A.; Bailey, C.C.; Boedeker, E.C.; Kaper, J.B.; Reeves, P.
 Proc. Natl. Acad. Sci. U.S.A. 95, 3134-3139, 1998
 A:Title: A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains
 A:Reference number: Z16672; MUID:98169509; PMID:9501228
 A:Accession: T09437
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-506 <KAR>
 A:Cross-references: UNIPROT:P23240; UNIPARC:UPI0000000A86; EMBL:AF034434; NID:g3004923;
 A:Experimental source: strain N16961
 C:Genetics:
 A:Gene: aldA
 A>Note: part of the pathogenicity island (VPI); associated with epidemic and pandemic strains
 C:Function:
 A:Description: catalyzes the oxidation of aldehydes
 A:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology
 C:Keywords: NAD; oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 506;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
 |||||
 DB 261 IELGGK 266

RESULT 773
 A99458
 C:Species: *Sulfolobus solfataricus*
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
 C:Accession: A99458
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, H.
 arrest, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: *Sulfolobus solfataricus* complete genome.
 A:Reference number: A99139
 A:Accession: A99458
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-507 <KUR>
 A:Cross-references: UNIPROT:Q97V27; UNIPARC:UPI0000064871; GB:AE006641; NID:g13816159; E
 C:Genetics:
 A:Gene: alkK-3
 C:Superfamily: probable acyl-CoA ligase medium chain; acetate-CoA ligase homology

Query Match 5.1%; Score 6; DB 2; Length 507;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||
 DB 84 LLGGVL 89

RESULT 774
 T05156
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C:Accession: T05156
 R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkee, W.; Stiekema, W.; Bancroft, I.; New
 submitted to the Protein Sequence Database, August 1998
 A:Reference number: Z15400
 A:Accession: T05156
 A:Molecule type: DNA
 A:Residues: 1-508 <BEV>
 A:Cross-references: UNIPROT:O65413; UNIPARC:UPI00000A1161; EMBL:AL022603

A:Experimental source: cultivar Columbia; BAC clone F18E5
 C:Genetics:
 A:Map position: 4
 A:Introns: 45/1; 152/3; 360/3
 A:Notes: F18E5.100
 C:Superfamily: glucose transport protein
 C:Keywords: sugar transport; transmembrane protein

Query Match 5.1%; Score 6; DB 2; Length 508;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
 |||||
 DB 114 LLGGVL 119

RESULT 775
 T30547
 C:Species: *Pneumocystis carinii*
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 15-Jun-2001
 C:Accession: T30547
 R:Mei, Q.; Turner, R.E.; Sorial, V.; Klivington, D.; Angus, C.W.; Kovacs, J.A.
 Infect. Immun. 66, 4268-4273, 1998
 A:Title: Characterization of major surface glycoprotein genes of human *Pneumocystis car*
 A:Reference number: Z17905; MUID:98380374; PMID:9712777
 A:Accession: T30547
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-508 <MEI>
 A:Cross-references: UNIPARC:UPI000006A08C; EMBL:AF038556; NID:g3560524; PID:g3560527; P
 A:Experimental source: f.sp. hominis
 C:Genetics:
 A:Gene: msg4
 C:Superfamily: *Pneumocystis carinii* major surface glycoprotein MSG100

Query Match 5.1%; Score 6; DB 2; Length 508;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
 |||||
 DB 246 IELGGK 251

RESULT 776
 A70580
 C:Species: *Mycobacterium tuberculosis* (strain H37Rv)
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
 C:Accession: A70580
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon,
 Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: A70580
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-510 <COL>
 A:Cross-references: UNIPROT:O06220; UNIPARC:UPI000012F9BB; GB:Z95388; GB:AL123456; NID:
 A:Experimental source: strain H37Rv
 C:Genetics:
 A:Gene: murF
 C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 5.1%; Score 6; DB 2; Length 510;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
 Db 109 VLAALA 114

RESULT 777
 B87258
 hypothetical protein CC0075 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 A:Accession: B87258
 R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Smolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A>Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: B87258
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-511 <STO>
 A:Cross-references: UNIPROT:Q9ABZ5; UNIPARC:UPI00000C6F0B; GB:AE005673; NID:g13421172; B
 C:Genetics:
 A:Gene: CC0075

Query Match 5.1%; Score 6; DB 2; Length 511;
 Best Local Similarity 100.0%; Pred. No. 7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 Db 321 LAALAA 326

RESULT 778
 D81701
 GMP synthase TC0442 [imported] - Chlamydia muridarum (strain Nigg)
 C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 A:Accession: D81701
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255; PMID:10684935
 A:Accession: D81701
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-512 <TET>
 A:Cross-references: UNIPROT:Q9PKM3; UNIPARC:UPI0000057916; GB:AE002313; GB:AE002160; NID
 A:Experimental source: strain Nigg (MoPn)
 C:Genetics:
 A:Gene: TC0442
 C:Superfamily: GMP synthase (glutamine-hydrolyzing); trpG homology

Query Match 5.1%; Score 6; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
 Db 375 GKVLGL 380

RESULT 779
 B86445
 unknown protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 A:Accession: B86445
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: B86445
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-512 <STO>
 A:Cross-references: UNIPROT:Q9FVQ4; UNIPARC:UPI00000A1AAD; GB:AE005172; NID:g10801363; F
 C:Genetics:
 A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
 Db 305 VLAALA 310

RESULT 780
 T21887
 hypothetical protein F36H2.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 A:Accession: T21887
 R:Steward, C.
 submitted to the EMBL Data Library, October 1996
 A:Reference number: Z19483
 A:Accession: T21887
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-513 <WIL>
 A:Cross-references: UNIPROT:Q93690; UNIPARC:UPI000017B98C; EMBL:Z81078; PIDN:CAB03075.1;
 A:Experimental source: clone F36H2
 C:Genetics:
 A:Gene: CESP:F36H2.2
 A:Map position: 1
 A:Introns: 28/3; 56/1; 93/2; 163/3; 277/3; 314/1; 341/3; 384/2; 407/2; 458/2; 480/3

Query Match 5.1%; Score 6; DB 2; Length 513;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
 Db 175 GVLAAL 180

RESULT 781
 AF3524
 hypothetical protein BMEI10120 [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
 A:Accession: AF3524
 R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A>Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AF3524
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-515 <KUR>
 A:Cross-references: UNIPROT:Q8YDQ5; UNIPARC:UPI0000058367; GB:AE008918; PIDN:AAL53361.1;
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI10120

A:Map position: II

Query Match 5.1%; Score 6; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 VLLGG 19
|||||
Db 32 VLLGG 37

RESULT 782

E86270

Hypothetical protein P21P23.15 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: E86270

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.E.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86270

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-517 <STO>

A:Cross-references: UNIPROT:Q9LMX7; UNIPARC:UPI00000AA7AB; GB:AE005172; NID:g8920576; P

C:Genetics:

A:Map position: 1

C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C:Keywords: heme; iron; metalloprotein

F:459/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 517;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 72 VLAALA 77

RESULT 783

H87022

Hypothetical protein murf [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 05-Oct-2004

C:Accession: H87022

R:Colo, S.T.; Eigmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H

R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

eam, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S

A:Title: Massive gene decay in the leprosy bacillus.

A:Reference number: A86909; MUID:21128732; PMID:11234002

A:Accession: H87022

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-517 <STO>

A:Cross-references: UNIPROT:O69556; UNIPARC:UPI000012P9BD; GB:AL450380; NID:gl3092980; F

C:Genetics:

A:Gene: murf

C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match 5.1%; Score 6; DB 2; Length 517;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 111 VLAALA 116

RESULT 784

AF2370

serine/threonine kinase [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AF2370

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriq

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AF2370

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-520 <KUR>

A:Cross-references: UNIPROT:Q8YNP6; UNIPARC:UPI00000CEB62; GB:BA000019; PIDN:BA076217.1

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: all4518

Query Match 5.1%; Score 6; DB 2; Length 520;

Best Local Similarity 100.0%; Pred. No. 7.1e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
Db 34 VLLGGV 39

RESULT 785

TVFVMT

protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus (strain PA101T)

C:Species: Rous sarcoma virus

C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 05-Oct-2004

C:Accession: A42994

R;Dezelee, P.; Barnier, J.V.; Hampe, A.; Laugier, D.; Marx, M.; Galibert, F.; Calothy,

Virology 189, 556-567, 1992

A:Title: Small deletion in v-src SH3 domain of a transformation defective mutant of Rou

A:Reference number: A42994; MUID:92351554; PMID:1322589

A:Accession: A42994

A:Molecule type: DNA

A:Residues: 1-523 <DEZ>

A:Cross-references: UNIPROT:P31693; UNIPARC:UPI0000135F2B; GB:M84475

C:Genetics:

A:Gene: src

C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology,

C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; on;

F:145-242/Domain: SH3 homology <SH2>

F:145-242/Domain: SH2 homology <SH2>

F:262-520/Domain: protein kinase homology <KIN>

F:270-278/Region: protein kinase ATP-binding motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F:292/Active site: Lys #status predicted

F:413/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status experim

Query Match 5.1%; Score 6; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 7.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
|||||
Db 180 AYCLSV 185

RESULT 786

C70717

probable purH protein - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: C70717
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70717
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-523 <COL>
A:Cross-references: UNIPROT:P71553; UNIPARC:UPI0000132ADA; GB:Z79700; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: purH
C:Superfamily: purH bifunctional enzyme

Query Match 5.1%; Score 6; DB 2; Length 523;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
|||||
Db 159 GVLAAL 164

RESULT 787
S15619
L2 protein - human papillomavirus type 2a
C:Species: human papillomavirus type 2a
A:Note: host Homo sapiens (man)
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: S15619
R:Hirsch-Behnam, A.; Delius, H.; de Villiers, E.M.
A:Title: A comparative sequence analysis of two human papillomavirus (HPV) types 2a and Virus Res. 18, 81-98, 1990
A:Reference number: S15614; MUID:91188699; PMID:1964523
A:Accession: S15619
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-524 <HIR>
A:Cross-references: UNIPROT:P25487; UNIPARC:UPI0000138900; EMBL:X55964
C:Superfamily: papillomavirus L2 protein
C:Keywords: late protein

Query Match 5.1%; Score 6; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 EVTTST 13
|||||
Db 185 EVTTST 190

RESULT 788
S55097
probable membrane protein YMR215w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YMR261.09
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S55097
R:Dedman, K.; Brown, D.; Bowman, S.
submitted to the EMBL Data Library, June 1995
A:Reference number: S55089
A:Accession: S55097
A:Molecule type: DNA
A:Residues: 1-524 <DED>
A:Cross-references: UNIPROT:Q03655; UNIPARC:UPI000012B084; EMBL:Z49809; NID:g854459; PID
A:Experimental source: strain AB372

C:Genetics:
A:Gene: MIPS:YMR215w
A:Cross-references: SGD:S0004828
A:Map position: 13R
C:Superfamily: glycopospholipid-anchored surface glycoprotein GAS1
C:Keywords: transmembrane protein
P:7-23/Domain: transmembrane #status predicted <TMM>

Query Match 5.1%; Score 6; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|||||
Db 9 LAALAA 14

RESULT 789
OKPVVR
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus (strain H-19)
N:Alternate names: kinase-related transforming protein src
C:Species: Rous sarcoma virus
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-Oct-2004
C:Accession: S09609
R:Bodor, J.; Poliak, E.; Pichrtova, J.; Geryk, J.; Svoboda, J.
Nucleic Acids Res. 17, 8869, 1989
A:Title: Complete nucleotide sequence of LTR, v-src, LTR provirus H-19.
A:Reference number: S09609; MUID:90067864; PMID:2587228
A:Accession: S09609
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-526 <BOD>
A:Cross-references: UNIPROT:P25020; UNIPARC:UPI0000135P2A; EMBL:X15345; NID:g61706; PID
C:Genetics:
A:Gene: src
C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; onc
F:88-137/Domain: SH3 homology <SH3>
F:148-245/Domain: SH2 homology <SH2>
F:265-523/Domain: protein kinase homology <KIN>
F:273-281/Region: protein kinase ATP-binding motif
P:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:295/Active site: Lys #status predicted
F:416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
|||||
Db 183 AYCLSV 188

RESULT 790
TVRV60
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus
C:Species: Rous sarcoma virus
C>Date: 22-May-1981 #sequence_revision 17-Dec-1982 #text_change 05-Oct-2004
C:Accession: A38017; A00631; S02726; A38018
R:Czernilofsky, A.P.; Levinson, A.D.; Varmus, H.E.; Bishop, J.M.; Tischler, E.; Goodman, Nature 301, 736-738, 1983
A:Title: Corrections to the nucleotide sequence of the src gene of Rous sarcoma virus.
A:Reference number: A38017; MUID:83141780; PMID:6298633
A:Accession: A38017
A:Molecule type: DNA
A:Residues: 1-526 <CZB>
A:Cross-references: UNIPROT:P00524; UNIPARC:UPI0000170DC3; GB:J29199; GB:J202018; GB:J020
A:Experimental source: strain Schmidt-Ruppin
R:Takeya, T.; Hanafusa, H.
Cell 32, 881-890, 1983
A:Title: Structure and sequence of the cellular gene homologous to the RSV src gene and
A:Reference number: A00630; MUID:83155664; PMID:6299580

A;Accession: A00631
A;Molecule type: DNA
A;Residues: 1-62, 'D', '64-95', 'T', '97-123', 'V', '125-300', 'N', '302-526 <TAK>
A;Cross-references: UNIPARC:UPI0000172582
A;Experimental source: strain Schmidt-Ruppin
R;Barnier, J.V.; Dezeles, P.; Marx, M.; Calothy, G.
Nucleic Acids Res. 17, 1252, 1989
A;Title: Nucleotide sequence of the src gene of the Schmidt-Ruppin strain of Rous Sarcoma virus
A;Reference number: S02726; MUID:89160256; PMID:2537953
A;Accession: S02726
A;Molecule type: DNA
A;Residues: 1-9, 'G', '11-62', 'D', '64-123', 'V', '125-319', 'K', '321-495', 'S', '497-526 <BAR>
A;Cross-references: UNIPARC:UPI0000135F2C; EMBL:X13745; NID:g61908; PIDN:CAA32012.1; PIDN:CAA32012.1; PIDN:CAA32012.1
R;Takeya, T.; Feldman, R.A.; Hanafusa, H.
J. Virol. 44, 1-11, 1982
A;Title: DNA sequence of the viral and cellular src gene of chickens: I. Complete nucleotide sequence of the viral src gene
A;Reference number: A38018; MUID:83059858; PMID:6292477
A;Accession: A38018
A;Molecule type: DNA
A;Residues: 1-15, 'C', '17-94', 'RT', '97-116', 'D', '118-337', 'T', '339-526 <TA2>
A;Cross-references: UNIPARC:UPI0000135F24; GB:K00928; NID:g210187; PIDN:AAA42565.1; PIDN:AAA42565.1; PIDN:AAA42565.1
A;Experimental source: strain rASV1441
R;Neil, J.C.; Ghysdael, J.; Vogt, P.K.; Smart, J.E.
Nature 291, 675-677, 1981
A;Title: Homologous tyrosine phosphorylation sites in transformation-specific gene products of Rous sarcoma virus
A;Reference number: A38019; MUID:81220979; PMID:6264320
A;Contents: annotation; phosphorylation site
C;Comment: The sequence from the Schmidt-Ruppin strain is shown.
C;Genetics:
A;Gene: src
C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; onco
F;148-245/Domain: SH3 homology <SH3>
F;188-137/Domain: SH3 homology <SH3>
F;265-523/Domain: protein kinase homology <KIN>
F;273-281/Region: protein kinase ATP-binding motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;295/Active site: Lys #status predicted
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status experiment

Query Match 5.1%; Score 6; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 AYCLSV 31
Db 183 AYCLSV 188

RESULT 791
TFVPR
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus (strain Prague C)
C;Species: Rous sarcoma virus
C;Date: 01-Sep-1981 #sequence_revision 17-Dec-1982 #text_change 05-Oct-2004
C;Accession: A00632
R;Schwartz, D.; Tizard, R.; Gilbert, W.
submitted to the Nucleic Acid Sequence Database, September 1982
A;Reference number: A00632
A;Accession: A00632
A;Molecule type: genomic RNA
A;Residues: 1-526 <SCH>
A;Cross-references: UNIPROT:P00526; UNIPROT:O92806; UNIPARC:UPI000002BA63
A;Note: as a result of base variations, residues 242 and 288 may be replaced by Thr and R;Neil, J.C.; Ghysdael, J.; Vogt, P.K.; Smart, J.E.
Nature 291, 675-677, 1981
A;Title: Homologous tyrosine phosphorylation sites in transformation-specific gene products of Rous sarcoma virus
A;Reference number: A38019; MUID:81220979; PMID:6264320
A;Contents: annotation; phosphorylation site
C;Genetics:
A;Gene: src
C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; onco
F;148-137/Domain: SH3 homology <SH3>
F;265-523/Domain: protein kinase homology <KIN>
F;273-281/Region: protein kinase ATP-binding motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;295/Active site: Lys #status predicted
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status experiment

F;148-245/Domain: SH2 homology <SH2>
F;265-523/Domain: protein kinase homology <KIN>
F;273-281/Region: protein kinase ATP-binding motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;295/Active site: Lys #status predicted
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status experiment

Query Match 5.1%; Score 6; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 AYCLSV 31
Db 183 AYCLSV 188

RESULT 792
S15582
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus (strain Prague A)
C;Species: Rous sarcoma virus
A;Variety: strain Prague A
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 05-Oct-2004
C;Accession: S15582; S09665
R;Liu, Z.; Hackett, P.B.
Nucleic Acids Res. 17, 3986, 1989
A;Title: Sequence variation of the Rous sarcoma virus PrA src gene.
A;Reference number: S15582; MUID:89282411; PMID:2543959
A;Accession: S15582
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-526 <LIU>
A;Cross-references: UNIPROT:Q64994; UNIPROT:O92806; UNIPROT:Q60567; UNIPROT:Q07461; UNIPROT:Q07461; UNIPROT:Q07461
A;Experimental source: strain Prague A
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1989
R;Fincham, V.J.; Wyke, J.A.
J. Virol. 58, 694-699, 1986
A;Title: Localization of temperature-sensitive transformation mutations and back mutations in the src gene of Rous sarcoma virus
A;Reference number: S09665; MUID:86200422; PMID:3009882
A;Accession: S09665
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 231-241, 'TH', '244-287', 'G', '289-463', 'P', '465-501', 'N', '503-526 <FIN>
A;Cross-references: UNIPARC:UPI00001755F1
C;Genetics:
A;Gene: src
C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; onco
F;148-137/Domain: SH3 homology <SH3>
F;148-245/Domain: SH2 homology <SH2>
F;265-523/Domain: protein kinase homology <KIN>
F;273-281/Region: protein kinase ATP-binding motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;295/Active site: Lys #status predicted
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 AYCLSV 31
Db 183 AYCLSV 188

RESULT 793
S20808
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus
C;Species: Rous sarcoma virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-Oct-2004
C;Accession: S20808; S32774
R;Bodor, J.; Rozkot, F.; Svoboda, J.
submitted to the EMBL Data Library, May 1990

A;Description: Sequence organization of the adjacent chromosomal flanks the LTR.
A;Reference number: S20808
A;Accession: S20808
A;Molecule type: DNA
A;Residues: 1-526 <BOD>
A;Cross-references: UNIPROT:Q60567; UNIPARC:UPI00001068B2; EMBL:X52822; NID:g49656; PIDN:R188-137/Domain: SH3 homology <SH3>
A;Experimental source: Mesocricetus auratus (golden hamster) provirus
C;Genetics:
A;Gene: src
A;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; transmembrane
F;148-245/Domain: SH3 homology <SH3>
F;265-523/Domain: protein kinase homology <KIN>
F;273-281/Region: protein kinase ATP-binding motif
F;295/Active site: Lys #status predicted
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
|||||
Db 183 AYCLSV 188

RESULT 794
S26420
A;Description: protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus
C;Species: Rous sarcoma virus
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Oct-2004
A;Accession: S26420; S20676
R;Kashuba, V.I.; Rynditch, A.V.; Dostalova, V.; Hlozanek, I.; Zubak, S.V.; Kavean, V.M. submitted to the EMBL Data Library, September 1992
A;Description: Molecular cloning and DNA sequence analysis of duck-adapted variant of Rous sarcoma virus
A;Reference number: S26417
A;Accession: S26420
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-526 <KAS>
A;Cross-references: UNIPROT:Q07461; UNIPARC:UPI000010512B; EMBL:X68524; NID:g61903; PIDN:R188-137/Domain: SH3 homology <SH3>
A;Reference number: S20676
A;Accession: S20676
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-526 <KA2>
A;Cross-references: UNIPARC:UPI000010512B; EMBL:X51861; NID:g61896; PIDN:CAA36154.1; PIDN:R188-137/Domain: SH3 homology <SH3>
C;Genetics:
A;Gene: src
A;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phosphatase
F;148-245/Domain: SH3 homology <SH3>
F;265-523/Domain: protein kinase homology <KIN>
F;273-281/Region: protein kinase ATP-binding motif
F;295/Active site: myristylated amino end (Gly) (in mature form) #status predicted
F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
|||||
Db 183 AYCLSV 188

RESULT 795
B70859

hypothetical protein Rv3031 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
A;Accession: B70859
R;Cole, S.T.; Davies, R.; Parkhill, J.; Garnier, T.; Churche, C.; Harris, D.; Gordon, S.; Connor, R.; Bross, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A;Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: B70859
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-526 <COL>
A;Cross-references: UNIPROT:O53278; UNIPARC:UPI00000D101E; GB:AL021287; GB:AL123456; NID:R188-137/Domain: SH3 homology <SH3>
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: Rv3031
C;Superfamily: Pyrococcus horikoshii hypothetical protein PH1386

Query Match 5.1%; Score 6; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 7.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
|||||
Db 50 VLAALA 55

RESULT 796
H82801
A;Description: conserved hypothetical protein XF0470 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
A;Accession: H82801
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: H82801
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-530 <SIM>
A;Cross-references: UNIPROT:Q9PG32; UNIPARC:UPI00000C240B; GB:AE003897; GB:AE003849; NID:R188-137/Domain: SH3 homology <SH3>
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Carrer, H.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Klieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.; A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir M.; Tuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; V A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0470

Query Match 5.1%; Score 6; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 7.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 AVIEPI 103
|||||
Db 231 AVIEPI 236

RESULT 797

TS1922
hypothetical protein B23111.340 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
R:Accession: T51922
R:Schulte, U.; Algn, V.; Hobeisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, August 2000
A:Reference number: 225958
A:Accession: T51922
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-531 <SCH>
A:Cross-references: UNIPARC:UPI000017B479; EMBL:AL391572; GSPDB:GN00116; NCSP:B23111.340
A:Experimental source: BAC clone B23111; strain OR74A
C:Genetics:
A:Gene: NCSP:B23111.340
A:Map position: 6
A:Introns: 42/3; 439/3; 505/2

Query Match 5.1%; Score 6; DB 2; Length 531;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 35 VLAALA 40

RESULT 798

B34104
protein-tyrosine kinase (EC 2.7.1.112) src 2 [similarity] - African clawed frog
N:Alternate names: Kinase-related transforming protein (src); kinase-related transformin
C:Species: Xenopus laevis (African clawed frog)
C>Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 05-Oct-2004
C:Accession: B34104; I51563
R:Steele, R.E.; Unger, T.F.; Mardis, M.J.; Fero, J.B.
J. Biol. Chem. 264, 10649-10653, 1989
A:Title: The two Xenopus laevis SRC genes are co-expressed and each produces functional
A:Reference number: A34104; MUID:89278134; PMID:2499582
A:Accession: B34104

A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-532 <STE>
A:Cross-references: UNIPROT:P13116; UNIPARC:UPI000017159F; GB:M23422; GB:J04822; NID:921
R:Steele, R.E.

Nucleic Acids Res. 13, 1747-1761, 1985
A:Title: Two divergent cellular src genes are expressed in Xenopus laevis.
A:Reference number: I51563; MUID:85215578; PMID:2987836
A:Accession: I51563
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 439-492 <ST2>
A:Cross-references: UNIPARC:UPI00001715A0; GB:M30858; NID:G214799; PIDN:AAA51644.1; PID:
C:Genetics:
A:Gene: src

A:Introns: 464/1
C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho
F:87-136/Domain: SH3 homology <SH3>
F:147-244/Domain: SH2 homology <SH2>
F:264-522/Domain: protein kinase homology <KIN>
F:272-280/Region: protein kinase ATP-binding motif
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:294/Active site: Lys #status predicted
F:415,526/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 5.1%; Score 6; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
|||||

DB 182 AYCLSV 187

RESULT 799

A34104
protein-tyrosine kinase (EC 2.7.1.112) src 1 [similarity] - African clawed frog
N:Alternate names: Kinase-related transforming protein (src); kinase-related transformi
C:Species: Xenopus laevis (African clawed frog)
C>Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 31-Dec-2004
C:Accession: A34104; I51564
R:Steele, R.E.; Unger, T.F.; Mardis, M.J.; Fero, J.B.
J. Biol. Chem. 264, 10649-10653, 1989
A:Title: The two Xenopus laevis SRC genes are co-expressed and each produces functional
A:Reference number: A34104; MUID:89278134; PMID:2499582
A:Accession: A34104

A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-532 <STE>
A:Cross-references: UNIPROT:Q91851; UNIPARC:UPI0000172581; GB:M24704; GB:J04822; NID:921
R:Steele, R.E.; Chosen, R.; Ral, B.B.A.; Winokur, S.T.; Unger, T.F.
Oncogene 7, 2345-2350, 1992

A:Title: Structural organization of a src gene from xenopus laevis.
A:Reference number: I51564; MUID:93064714; PMID:1437158
A:Accession: I51564

A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-113 <ST2>
A:Cross-references: UNIPARC:UPI00000FD97A; GB:M33646; NID:G214808; PIDN:AAA49963.1; PID:
C:Genetics:

A:Introns: 80/1
C:Superfamily: protein kinase homology; SH2 homology; SH3 homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; ph

F:87-136/Domain: SH3 homology <SH3>
F:147-244/Domain: SH2 homology <SH2>
F:264-522/Domain: protein kinase homology <KIN>
F:272-280/Region: protein kinase ATP-binding motif
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:294/Active site: Lys #status predicted
F:415,526/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pre

Query Match 5.1%; Score 6; DB 1; Length 532;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
|||||

DB 182 AYCLSV 187

RESULT 800

A57173
oculocutaneous albinism II-related protein P - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 17-Mar-1999
C:Accession: A57173
R:Lee, S.T.; Nicholls, R.D.; Jong, M.T.C.; Fukai, K.; Spritz, R.A.
Genomics 26, 354-363, 1995
A:Title: Organization and sequence of the human P gene and identification of a new fami
A:Reference number: A57173; MUID:95324928; PMID:7601462
A:Accession: A57173

A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-532 <LEE>
A:Cross-references: UNIPARC:UPI000017C2D7; GB:U19152
C:Genetics:
A:Gene: GDB:OCA2; P; D15S12

A:Map position: 15ql1.2-15ql2
Query Match 5.1%; Score 6; DB 2; Length 532;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 55 LAALAA 60

RESULT 801
TVCHS
protein-tyrosine kinase (EC 2.7.1.112) src - chicken
N/Alternate names: kinase-related transforming protein src
C/Species: Gallus gallus (Chicken)
C/Date: 19-Feb-1984 #sequence revision 07-Oct-1994 #text_change 05-Oct-2004
C/Accession: A00630; I50217; A41256; C35650; A32432
R/Takeya, T.; Hanafusa, H.
Cell 32, 881-890, 1983
A/Title: Structure and sequence of the cellular gene homologous to the RSV src gene and
A/Reference number: A00630; MUID:83155664; PMID:6299580
A/Accession: A00630
A/Molecule type: DNA
A/Residues: 1-500,'R',502-533 <TAK>
A/Cross-references: UNIPROT:P00523; UNIPARC:UPI000017257F; GB:J00844; NI
R/Takeya, T.; Hanafusa, H.
Cell 34, 319, 1983
A/Reference number: A30838
A/Contents: annotation; erratum, correct translation of residue 526
R/Takeya, T.; Hanafusa, H.
J. Virol. 44, 12-18, 1982
A/Title: DNA sequence of the viral and cellular src gene of chickens: II comparison of b
A/Reference number: I50217; MUID:83059861; PMID:6292480
A/Accession: I50217
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-7 <TA2>
A/Cross-references: UNIPARC:UPI000011E887; GB:J00908; NID:9211690; PIDN:AAA48732.1; PID:
R/Dorai, T.; Levy, J.B.; Kang, L.; Brugges, J.S.; Wang, L.H.
Mol. Cell. Biol. 11, 4165-4176, 1991
A/Title: Analysis of cDNAs of the proto-oncogene c-src: heterogeneity in 5' exons and po
A/Reference number: A41256; MUID:91304409; PMID:1712905
A/Accession: A41256
A/Molecule type: mRNA
A/Residues: 484-533 <DOR1>
A/Cross-references: UNIPARC:UPI0000171468; GB:S43579; NID:91679964; PIDN:AA19353.1; PID:
R/Dorai, T.; Wang, L.H.
Mol. Cell. Biol. 10, 4068-4079, 1990
A/Title: An alternative non-tyrosine protein kinase product of the c-src gene in chicken
A/Reference number: A35650; MUID:90318371; PMID:2115117
A/Accession: C35650
A/Residues: 1-182,'DPCIPLPCLC' <DOR2>
A/Cross-references: UNIPARC:UPI00000FD3A4; GB:M57290; NID:9212703; PIDN:AAA49078.1; PID:
A/Note: alternatively spliced mRNA exclusively replaces the long form in skeletal muscle
A/Note: this ORF appears not to be translated
R/Shenoy, S.; Choi, J.K.; Bagrodia, S.; Copeland, T.D.; Maller, J.L.; Shalloway, D.
Cell 57, 763-774, 1989
A/Title: Purified maturation promoting factor phosphorylates pp60(c-src) at the sites ph
A/Reference number: A32432; MUID:89249341; PMID:2470512
A/Accession: A32432
A/Molecule type: protein
A/Residues: 2-88 <SH2>
A/Cross-references: UNIPARC:UPI0000172580
A/Note: 34-Thr, 46-Thr, and 72-Ser are phosphorylated during mitosis
C/Genetics: src
A/Gene: src
C/Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C/Keywords: alternative splicing; ATP; autophosphorylation; blocked amino end; lipoprote
F/88-137/Domain: SH3 homology <SH3>
F/148-245/Domain: SH2 homology <SH2>
F/265-523/Domain: protein kinase homology <KIN>
F/273-281/Region: protein kinase ATP-binding motif
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F/12,48/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
F/17/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

F/34,46/Binding site: phosphate (Thr) (covalent) #status experimental
F/72/Binding site: phosphate (Ser) (covalent) #status experimental
F/295/Active site: Lys #status predicted
F/416,527/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 5.1%; Score 6; DB 1; Length 533;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
|||||
Db 183 AYCLSV 188

RESULT 802
D98224
dipeptide transport protein (AB036425) [imported] - Agrobacterium tumefaciens (strain C5)
C/Species: Agrobacterium tumefaciens
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C/Accession: D98224
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: D98224
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-534 <KUR>
A/Cross-references: UNIPROT:Q8U8H9; UNIPARC:UPI00000D244B; GB:AE007870; PIDN:AAK89318.1;
C/Genetics:
A/Gene: AGR_L1481
A/Map position: linear chromosome
C/Superfamily: dipeptide transport protein

Query Match 5.1%; Score 6; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EQAQVI 78
|||||
Db 486 EQAQVI 491

RESULT 803
AC3062
hypothetical protein dppA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C/Accession: AC3062
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
Sker, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AC3062
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-534 <KUR>
A/Cross-references: UNIPROT:Q8U8H9; UNIPARC:UPI00000D244B; GB:AE008689; PIDN:AA144913.1;
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: dppA
A/Map position: linear chromosome
C/Superfamily: dipeptide transport protein

Query Match 5.1%; Score 6; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EOAQVI 78
|||||
Db 486 EOAQVI 491

RESULT 804

T34455

hypothetical protein T19H12.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

A:Accession: T34455

R:Davidson, S.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T19H12.

A:Reference number: Z21528

A:Accession: T34455

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-534 <DAV>

A:Cross-references: UNIPROT:O01614; UNIPARC:UPI000008358A; EMBL:U97009; PIDN:AAC69032.1;

A:Experimental source: strain Bristol N2; clone T19H12

C:Genetics:

A:Gene: CESP:T19H12.9

A:Map position: 5

A:Introns: 51/3; 127/1; 157/1; 444/2

C:Superfamily: glucuronosyltransferase

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 534;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GKPAIV 49

|||||

Db 393 GKPAIV 398

RESULT 805

H83324

probable chemotaxis transducer PA2573 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

A:Accession: H83324

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: H83324

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-535 <STO>

A:Cross-references: UNIPROT:Q910R3; UNIPARC:UPI00000C5747; GB:AE004685; GB:AE004091; NID

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2573

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 535;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20

|||||

Db 19 VLLGGV 24

RESULT 806

C82433

methyl-accepting chemotaxis protein VCA0658 [imported] - Vibrio cholerae (strain N16961

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

A:Accession: C82433

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragoi, I.; Sellers,
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: C82433

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-536 <HEI>

A:Cross-references: UNIPROT:Q9KLT3; UNIPARC:UPI00000C35D8; GB:AE004395; GB:AE003853; N1

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0658

A:Map position: 2

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 536;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 QRATQQ 96

|||||

Db 415 QRATQQ 420

RESULT 807

A13078

conserved hypothetical protein Atu4253 [imported] - Agrobacterium tumefaciens (strain C

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

A:Accession: A13078

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: A82577; MUID:21608550; PMID:11743193

A:Accession: A13078

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-536 <KUR>

A:Cross-references: UNIPROT:Q8U846; UNIPARC:UPI00001648F7; GB:AE008689; PIDN:AAL45047.1;

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu4253

A:Map position: linear chromosome

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 536;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAYCL 29

|||||

Db 493 LAAYCL 498

RESULT 808

A29923

carboxylesterase (EC 3.1.1.1), TCCD-induced 60K microsomal - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 08-Dec-1988 #sequence_revision 08-Dec-1988 #text_change 09-Jul-2004

A:Accession: A29923; A29471

R:Korza, G.; Ozols, J.

J. Biol. Chem. 263, 3486-3495, 1988

A:Title: Complete covalent structure of 60-kDa esterase isolated from 2,3,7,8-tetrachlo

A:Reference number: A29923; MUID:88139431; PMID:3343253

A:Accession: A29923

A:Molecule type: protein

A:Residues: 1-539 <KOR>

A:Cross-references: UNIPROT:P12337; UNIPARC:UPI000012A1EF

R:Ozols, J.

J. Biol. Chem. 262, 15316-15321, 1987

A;Title: Isolation and characterization of a 60-kilodalton glycoprotein esterase from 14
A;Reference number: A29471; MUID:88033124; PMID:3667634
A;Accession: A29471
A;Molecule type: protein
A;Residues: 1-71;193-208;436-446;532-539 <OZO>
A;Cross-references: UNIPARC:UPI00001759F8; UNIPARC:UPI00001759F9; UNIPARC:UPI00001759FA;
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein
F;32-526/Domain: cholinesterase homology <CHS>
F;61.363/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;195,441/Active site: Ser, His #status experimental

Query Match 5.1%; Score 6; DB 2; Length 539;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLG 88
|||||
Db 12 KGKVLG 17

RESULT 809
G87407
oxidoreductase, GMC family CCl278 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
A;Accession: G87407
R;Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: G87407
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-540 <STO>
A;Cross-references: UNIPROT:Q9A8S3; UNIPARC:UPI00000C7322; GB:AE005673; NID:gl3422611; F
C;Genetics:
A;Gene: CCl278
C;Superfamily: alcohol oxidase

Query Match 5.1%; Score 6; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
|||||
Db 440 AALAA 445

RESULT 810
A31584
carboxylesterase (EC 3.1.1.1) precursor - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 21-May-1990 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
A;Accession: A31584
R;Long, R.M.; Satoh, H.; Martin, B.M.; Kimura, S.; Gonzalez, F.J.; Pohl, L.R.
Biochem. Biophys. Res. Commun. 156, 866-873, 1988
A;Title: Rat liver carboxylesterase: cDNA cloning, sequencing, and evidence for a multi
A;Reference number: A31584; MUID:89050119; PMID:2973315
A;Accession: A31584
A;Molecule type: mRNA
A;Residues: 1-540 <LON>
A;Cross-references: UNIPROT:P10959; UNIPARC:UPI00001708B2; GB:M20629; GB:X13587; NID:g20
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase; glycoprotein
F;1-3/Domain: signal sequence #status predicted <SIG>
F;10-540/Product: carboxylesterase #status predicted <MAT>
F;41-529/Domain: cholinesterase homology <CHE>
F;70,265,266,293,366,467/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;212,444/Active site: Ser, His #status predicted

Query Match 5.1%; Score 6; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLG 88
|||||
Db 21 KGKVLG 26

RESULT 811
G75386
probable 2-phosphoglycerate kinase - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
A;Accession: G75386
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-540 <WHI>
A;Cross-references: UNIPROT:Q9RU78; UNIPARC:UPI00000D3E73; GB:AE001995; GB:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1514
A;Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGV 20
|||||
Db 312 VLLGGV 317

RESULT 812
A43610
protein-tyrosine kinase (EC 2.7.1.112) src, neuronal [similarity] - mouse
N;Alternate names: Rouse sarcoma oncogene
C;Species: Mus musculus (house mouse)
C;Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 05-Oct-2004
A;Accession: A43610
R;Martinez, R.; Mathey-Prevot, B.; Bernards, A.; Baltimore, D.
Science 237, 411-415, 1987
A;Title: Neuronal pp60(c-src) contains a six-amino acid insertion relative to its non-ne
A;Reference number: A43610; MUID:87263406; PMID:2440106
A;Accession: A43610
A;Molecule type: mRNA
A;Residues: 1-541 <MAR>
A;Cross-references: UNIPROT:P05480; UNIPARC:UPI0000161D19; GB:M17031; NID:g201056; PIDN:
C;Comment: The neuronal c-src has an 6 residue insertion of RLNVNR within the amino-term
C;Genetics:
A;Gene: Src
A;Cross-references: MGI:98397
A;Map position: 2:91.0
C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho
F;90-145/Domain: SH3 homology <SH3>
F;156-253/Domain: SH2 homology <SH2>
F;273-531/Domain: protein kinase homology <KIN>
F;281-289/Region: protein kinase ATP-binding motif
F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F;303/Active site: Lys #status predicted
F;424,535/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pre

Query Match 5.1%; Score 6; DB 1; Length 541;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
 DB 191 AYCLSV 196

RESULT 813
 A82276
 aldehyde dehydrogenase VC0819 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 15-Mar-2004
 C:Accession: A82276
 R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
 chardson, D.; Emolaeva, M.D.; Vamathevan, J.; Baas, S.; Qin, H.; Dragoi, I.; Sellers, E.
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
 A:Reference number: A82035; MUID:20406833; PMID:10952301
 A:Accession: A82276
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-541 <HEI>
 A:Cross-references: UNIPARC:UPI0000164B65; GB:AE004167; GB:AE003852; NID:g9655268; PIDN:
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VC0819
 A:Map position: 1
 C:Superfamily: NAD-dependent aldehyde dehydrogenase; aldehyde dehydrogenase homology

Query Match 5.1%; Score 6; DB 2; Length 541;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGK 45
 DB 296 IELGK 301

RESULT 814
 H83445
 Probable chemotaxis transducer PA1608 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: H83445
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B.
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: H83445
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-541 <STO>
 A:Cross-references: UNIPROT:Q913B3; UNIPARC:UPI00000C5406; GB:AE004588; GB:AE004091; NID
 A:Experimental source: strain PAO1
 C:Genetics:
 A:Gene: PA1608

Query Match 5.1%; Score 6; DB 2; Length 541;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 92 RATQQQ 97
 DB 157 RATQQQ 162

RESULT 815
 TWHUSC
 protein-tyrosine kinase (EC 2.7.1.112) src, neuronal - human
 C:Species: Homo sapiens (man)
 C>Date: 30-Jun-1989 #sequence_revision 07-Oct-1994 #text_change 05-Oct-2004

C:Accession: A26891; A61083; B61083; A23287; A28832; B34704
 R:Tanaka, A.; Gibbs, C.P.; Arthur, R.R.; Anderson, S.K.; Kung, H.J.; Fujita, D.J.
 Mol. Cell. Biol. 7, 1978-1983, 1987
 A:Title: DNA sequence encoding the amino-terminal region of the human c-src protein: im
 A:Reference number: A26891; MUID:87257903; PMID:3299057
 A:Accession: A26891
 A:Molecule type: mRNA
 A:Residues: 1-117;124-191 <TAN>
 A:Cross-references: UNIPROT:PI2931; UNIPARC:UPI0000172578; UNIPARC:UPI0000172579; GB:M1
 R:Pyper, J.M.; Bolen, J.B.
 J. Neurosci. Res. 24, 89-96, 1989
 A:Title: Neuron-specific splicing of C-SRC RNA in human brain.
 A:Reference number: A61083; MUID:90040822; PMID:2681803
 A:Accession: A61083
 A:Molecule type: mRNA
 A:Residues: 98-145 <PYP>
 A:Cross-references: UNIPARC:UPI000017257A
 A:Accession: B61083
 A:Molecule type: mRNA
 A:Residues: 98-117;124-145 <PY2>
 A:Cross-references: UNIPARC:UPI000017257A
 R:Anderson, S.K.; Gibbs, C.P.; Tanaka, A.; Kung, H.J.; Fujita, D.J.
 Mol. Cell. Biol. 5, 1122-1129, 1985
 A:Title: Human cellular src gene: Nucleotide sequence and derived amino acid sequence o
 A:Reference number: A23287; MUID:85213483; PMID:2582238
 A:Accession: A23287
 A:Molecule type: mRNA
 A:Residues: 192-542 <AND>
 A:Cross-references: UNIPARC:UPI000016B068; GB:X02647; NID:g36588; PIDN:CAA26485.1; PID:
 R:Parker, R.C.; Mardon, G.; Lebo, R.V.; Varmus, H.E.; Bishop, J.M.
 Mol. Cell. Biol. 5, 831-838, 1985
 A:Title: Isolation of duplicated human c-src genes located on chromosomes 1 and 20.
 A:Reference number: A28832; MUID:85187981; PMID:2581127
 A:Accession: A28832
 A:Molecule type: mRNA
 A:Residues: 382-542 <PAR>
 A:Cross-references: UNIPARC:UPI000017257D
 R:Pyper, J.M.; Bolen, J.B.
 Mol. Cell. Biol. 10, 2035-2040, 1990
 A:Title: Identification of a novel neuronal C-SRC exon expressed in human brain.
 A:Reference number: A34704; MUID:90220588; PMID:1691439
 A:Accession: B34704
 A:Molecule type: mRNA
 A:Residues: 118-123 <PY3>
 A:Cross-references: UNIPARC:UPI000017257E
 C:Genetics:
 A:Gene: GDB:SRC
 A:Cross-references: GDB:120750; OMIM:190090
 A:Map position: 20q11.2-20q11.2
 A:Introns: 84/1; 117/2; 123/2; 156/2; 191/1; 241/1; 293/1; 353/1; 378/3; 430/1; 474/1
 C:Function:
 A:Description: catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
 C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; ph
 F:1-542/Product: protein-tyrosine kinase src, neuronal #status predicted <MAT>
 F:1-117,124-542/Product: protein-tyrosine kinase src, short form #status predicted <MA2>
 F:91-146/Domain: SH3 homology <SH3>
 F:157-254/Domain: SH2 homology <SH2>
 F:274-532/Domain: protein kinase homology <KIN>
 F:282-290/Region: protein kinase ATP-binding motif
 F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F:304/Active site: Lys #status predicted
 F:425,536/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pre

Query Match 5.1%; Score 6; DB 1; Length 542;
 Best Local Similarity 100.0%; Pred. No. 7.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
 DB 192 AYCLSV 197

RESULT 816

F64871
oligopeptide-binding protein precursor - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: F64871, A36263
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: F64871
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-543 <BLAT>
A:CROSS-references: UNIPROT:P23843; UNIPARC:UPI0000130D72; GB:AE000222; GB:U00096; NID:G
A:Experimental source: strain K-12, substrain MG1655
R:Kashiwagi, K.; Yamaguchi, Y.; Sakai, Y.; Kobayashi, H.; Igaraashi, K.
J. Biol. Chem. 265, 8387-8391, 1990
A:Title: Identification of the polyamine-induced protein as a periplasmic oligopeptide b
A:Reference number: A36263; MUID:90256749; PMID:2187863
A:Accession: A36263
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-270, 'Y', 272-313, 'LW', 316-486, 'HG', 489-543 <KAS>
A:CROSS-references: UNIPARC:UPI000016F363; GB:J055433; NID:G147013; PIDN:AAA21302.1; PID:
C:Genetics:
A:Gene: oppA
C:Superfamily: dipeptide transport protein
C:Keywords: binding protein-dependent transport system; oligopeptide transport; periplas
F:1-26/Domain: signal sequence #status predicted <Sig>
F:27-543/Product: oligopeptide-binding protein #status predicted <MAT>

Query Match 5.1%; Score 6; DB 1; Length 543;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
Db 13 GVLAAL 18

RESULT 817
F85704
hypothetical protein oppA [imported] - Escherichia coli (strain O157:H7, substrain EDL93
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: F85704
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
filler, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: F85704
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <STO>
A:CROSS-references: UNIPROT:O8XDA4; UNIPARC:UPI0000165793; GB:AE005174; NID:G12514962; E
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: oppA
C:Superfamily: dipeptide transport protein

G90846
hypothetical protein ECs1743 [imported] - Escherichia coli (strain O157:H7, substrain R1)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: G90846
R:Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic reference
A:Reference number: A39629; MUID:21156231; PMID:11258796
A:Accession: G90846
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-543 <HAY>
A:Cross-references: UNIPROT:Q8XDA4; UNIPARC:UPI00000D03A5; GB:BA000007; PIDN:BA035166.1;
A:Experimental source: strain O157:H7, substrain R1MD 0509952
C:Genetics:
A:Gene: ECs1743
C:Superfamily: dipeptide transport protein

	Query Match	5.1%	Score 6;	DB 2;	Length 543;
	Best Local Similarity	100.0%;	Pred. No. 7.4e+02;		
	Matches	6;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	19	GVLAAL 24			
Db	13	GVLAAL 18			

RESULT 819
S52313
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus
C:Species: Rous sarcoma virus
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
C:Accession: S52313
R:Tatossyan, A.; Yatsula, B.; Shtutman, M.; Moinova, E.; Kaverina, I.; Musatkina, E.; Lee
submitted to the EMBL Data Library, January 1995
A:Description: Two new isoforms of v-src oncogene isolated from low and high metastatic
A:Reference number: S52313
A:Accession: S52313
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-545 <TAT>
A:Cross-references: UNIPROT:Q86362; UNIPARC:UPI0000105D06; EMBL:X84074; NID:G663083; PID
C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pho
P108-157/Domain: SH3 homology <SH3>
P168-265/Domain: SH2 homology <SH2>
P1285-543/Domain: protein kinase homology <KIN>
P1293-301/Region: protein kinase ATP-binding motif
P12/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
P1315/Active site: Lys #status predicted
P1436/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

	Query Match	5.1%	Score 6;	DB 2;	Length 545;
	Best Local Similarity	100.0%;	Pred. No. 7.4e+02;		
	Matches	6;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	26	AYCLSV 31			
Db	203	AYCLSV 208			

RESULT 820
S52314
protein-tyrosine kinase (EC 2.7.1.112) src - Rous sarcoma virus
C:Species: Rous sarcoma virus
C:Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
C:Accession: S52314
R:Tatossyan, A.; Yatsula, B.; Shtutman, M.; Moinova, E.; Kaverina, I.; Musatkina, E.; Lee
submitted to the EMBL Data Library, January 1995
A:Description: Two new isoforms of v-src oncogene isolated from low and high metastatic
A:Reference number: S52313

A:Accession: S52314
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <TAT>
A:Cross-references: UNIPROT:O86363; UNIPARC:UPI0000106213; EMBL:X84073; NID:g663085; PIDN:U22894
C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; pld
F:108-157/Domain: SH3 homology <SH3>
F:168-265/Domain: SH2 homology <SH2>
F:285-543/Domain: protein kinase homology <KIN>
F:293-301/Region: protein kinase ATP-binding motif
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:315/Active site: Lys #status predicted
F:436/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSV 31
|||
Db 203 AYCLSV 208

RESULT 821
T30269
hypothetical protein aglX - Streptomyces lividans
C:Species: Streptomyces lividans
C:Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 09-Jul-2004
C:Accession: T30269
R:Volff, J.N.; Eichenseer, C.; Viell, P.; Pendl, W.; Altenbuchner, J.
submitted to the EMBL Data Library, July 1998
A:Reference number: Z20798
A:Accession: T30269
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-547 <VOL>
A:Cross-references: UNIPROT:O86875; UNIPARC:UPI000008A0F9; EMBL:U22894; PIDN:AAC46451.1
C:Genetics:
A:Note: aglX
A:Note: amplifiable element AUD1

Query Match 5.1%; Score 6; DB 2; Length 547;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||
Db 451 VLAALA 456

RESULT 822
B84306
hypothetical protein Vngl525c [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: B84306
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laskey, S.; Leibauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: B84306
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <STO>
A:Cross-references: UNIPROT:Q9HPQ3; UNIPARC:UPI00000638EB; GB:AE004437; NID:g10581014; F
C:Genetics:
A:Gene: VNG1525C

Query Match 5.1%; Score 6; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||
Db 231 GVLAAL 236

RESULT 823
A75357
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: A75357
R:White, O.; Eisen, J.A.; Heideberg, J.P.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75357
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <WHI>
A:Cross-references: UNIPROT:Q9RTK8; UNIPARC:UPI00000C19B8; GB:AE002017; GB:AE000513; NI
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1752
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||
Db 28 VLLGGV 33

RESULT 824
H70788
probable peptidetransport system ABC-transporter ATP-binding protein - Mycobacterium tu
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
C:Accession: H70788
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: H70788
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-548 <COL>
A:Cross-references: UNIPROT:O69631; UNIPARC:UPI00000D5E42; GB:AL022121; GB:AL123456; NI
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: dppD
C:Keywords: ATP
F:307-508/Domain: ATP-binding cassette homology <ABC2>

Query Match 5.1%; Score 6; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||
Db 337 VLGLLQ 342

RESULT 825

JX0054
carboxylesterase (EC 3.1.1.1) E1 precursor, minor form - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 21-Aug-1998
C:Accession: JX0054
R:Takagi, Y.; Morohashi, K.; Kawabata, S.; Go, M.; Omura, T.
J. Biochem. 104, 801-806, 1998
A:Title: Molecular cloning and nucleotide sequence of cDNA of microsomal carboxylesterase
A:Reference number: JX0054; MUID:89174514; PMID:3235453
A:Accession: JX0054
A:Molecule type: mRNA
A:Residues: 1-549 <TAK>
A:Cross-references: UNIPARC:UPI00001758FC
A:Experimental source: liver
C:Superfamily: cholinesterase, cholinesterase homology
C:Keywords: carboxylic ester hydrolase; glycoprotein; microsome
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-549/Product: carboxylesterase E1 #status predicted <MAT>
F:50-538/Domain: cholinesterase homology <CHE>
F:79,274,302,375,476/Binding site: carboxylate (Asn) (covalent) #status predicted
F:721,453/Active site: Ser, His #status predicted

Query Match 5.1%; Score 6; DB 2; Length 549;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 GKGVLG 88
|||||
DB 30 GKGVLG 35

RESULT 826
H98207
hypothetical protein AGR_L1216 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: H98207
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: H98207
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-551 <KUR>
A:Cross-references: UNIPROT:Q8U846; UNIPARC:UPI0000D24CF; GB:AE007870; PIDN:AAK89186.1;
C:Genetics:
A:Gene: AGR_L1216
A:Map position: linear chromosome

Query Match 5.1%; Score 6; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 LAAAYCL 29
|||||
DB 508 LAAAYCL 513

RESULT 827
A11829
hypothetical protein all0185 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: A11829
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: A11829

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-554 <KUR>
A:Cross-references: UNIPROT:Q8Z0B3; UNIPARC:UPI00000CD6D; GB:BA000019; PIDN:BA077709.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0185
C:Superfamily: mvn protein

Query Match 5.1%; Score 6; DB 2; Length 554;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
DB 82 VLLGGV 87

RESULT 828
T06491
beta-fructofuranosidase (EC 3.2.1.26) - garden pea
N:Alternate names: invertase
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06491
R:Zhang, L.
submitted to the EMBL Data Library, January 1996
A:Reference number: Z15715
A:Accession: T06491
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-555 <ZHA>
A:Cross-references: UNIPROT:Q43089; UNIPARC:UPI000012D70F; EMBL:X85327; NID:g1160487; PI
A:Experimental source: cv. Little Marvel
C:Genetics:
A:Note: bfruct1
C:Superfamily: beta-fructofuranosidase
C:Keywords: glycosidase; hydrolase

Query Match 5.1%; Score 6; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GKGPAI 48
|||||
DB 129 GKGPAI 134

RESULT 829
F72555
probable molybdenum transport system permease protein APE1730 - Aeropyrum pernix (strain
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F72555
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72555
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-555 <KAW>
A:Cross-references: UNIPROT:Q9YB67; UNIPARC:UPI000005E077; DBJ:AP000062; NID:G5105244;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1730

Query Match 5.1%; Score 6; DB 2; Length 555;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21

Db 472 LLGGVL 477
|||||
RESULT 830
AG3432
chloride channel protein [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AG3432
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AG3432
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-555 <KUR>
A:Cross-references: UNIPROT:Q8YFSS; UNIPARC:UPI000005808B; GB:AE008917; PIDN:AAU52626.1
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11445
A:Map position: I
Query Match 5.1%; Score 6; DB 2; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 KVLGLL 90
|||||
Db 526 KVLGLL 531
protein-tyrosine kinase (EC 2.7.1.112) src - avian sarcoma virus S2
C:Species: avian sarcoma virus S2
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 05-Oct-2004
C:Accession: B25375
R:Ikawa, S.; Hagino-Yamagishi, K.; Kawai, S.; Yamamoto, T.; Toyoshima, K.
Mol. Cell. Biol. 6, 2420-2428, 1986
A:Title: Activation of the cellular src gene by transducing retrovirus.
A:Reference number: A25375; PMID:87064539; PMID:3097513
A:Accession: B25375
A:Molecule type: DNA
A:Residues: 1-557 <IKA>
A:Cross-references: UNIPROT:PI4085; UNIPARC:UPI0000135F26
C:Genetics:
A:Gene: src
C:Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; onc
F:88-137/Domain: SH3 homology <SH3>
F:148-245/Domain: SH2 homology <SH2>
F:265-523/Domain: protein kinase homology <KIN>
F:273-281/Region: protein kinase ATP-binding motif
F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F:295/Active site: Lys #status predicted
F:416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted
Query Match 5.1%; Score 6; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 26 AYCLSV 31
|||||
Db 183 AYCLSV 188
RESULT 832
B96106
yidB protein [similarity] - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B96106
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; PMID:21074935; PMID:11206551
A:Accession: B96106
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-557 <STO>
A:Cross-references: UNIPROT:Q8XEJ3; UNIPARC:UPI000000022D; GB:AE005174; NID:gl2519084;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yidB
C:Superfamily: Escherichia coli yidB protein
Query Match 5.1%; Score 6; DB 2; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GVLAAL 24
|||||
Db 142 GVLAAL 147
RESULT 833
S56342
yidB protein - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S56342; A65221
R:Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A:Title: Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92
A:Reference number: S56314; PMID:95334362; PMID:7610040
A:Accession: S56342
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-557 <BUR>
A:Cross-references: UNIPROT:Q8XEJ3; UNIPARC:UPI000000022D; EMBL:U14003; NID:gl263172; P
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; PMID:97426617; PMID:9278503
A:Accession: A65221
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-557 <BLAT>
A:Cross-references: UNIPARC:UPI000000022D; GB:AE000483; GB:U000096; NID:G2367351; PIDN:AF
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: yidB
C:Superfamily: Escherichia coli yidB protein
Query Match 5.1%; Score 6; DB 2; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GVLAAL 24
|||||
Db 142 GVLAAL 147
RESULT 834
H91265
hypothetical protein ECs5096 [imported] - Escherichia coli (strain O157:H7, substrain RI
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H91265
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic islands
 A;Reference number: A99629; MUID:21156231; PMID:11258796
 A;Accession: H91265
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-557 <HAY>
 A;Cross-references: UNIPROT:Q8XBJ3; UNIPARC:UPI0000165552; GB:BA0000007; PIDN:BA838519.1.7
 A;Experimental source: strain O157:H7, substrain RIMD 0509952
 C;Genetics:
 A;Gene: ECs5096
 C;Superfamily: *Escherichia coli* yidB protein

Query Match 5.1%; Score 6; DB 2; Length 557;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
 Db 142 GVLAAL 147
 |||||

RESULT 835
 B75477
 conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C;Species: *Deinococcus radiodurans*
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C;Accession: B75477
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A;Reference number: A75250; MUID:20036896; PMID:10567266
 A;Accession: B75477
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-559 <WHI>
 A;Cross-references: UNIPROT:Q9RW79; UNIPARC:UPI00000C180A; GB:AE001933; GB:AE000513; NID
 A;Experimental source: strain R1
 C;Genetics:
 A;Gene: DR0790
 A;Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 559;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 Db 155 LAALAA 160
 |||||

RESULT 836
 S71597
 carboxylesterase (EC 3.1.1.1) precursor, liver - rat
 N;Alternate names: hydrolase C
 C;Species: *Rattus norvegicus* (Norway rat)
 C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 13-Sep-1998
 C;Accession: S71597
 R;Yan, B.; Yang, D.; Parkinson, A.
 Arch. Biochem. Biophys. 317, 222-234, 1995
 A;Title: Cloning and expression of hydrolase C, a member of the rat carboxylesterase fam
 A;Reference number: S71597; MUID:9517656; PMID:7672788
 A;Accession: S71597
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-561 <YAN>
 A;Cross-references: UNIPARC:UPI00001758F5
 A;Experimental source: liver; endoplasmic reticulum
 C;Function:
 A;Description: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion

C;Superfamily: cholinesterase; cholinesterase homology
 C;Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-561/Product: carboxylesterase #status predicted <MAT>
 F;50-551/Domain: cholinesterase homology <CHE>
 F;558-561/Region: endoplasmic reticulum retention signal
 F;79,301/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;221,466/Active site: Ser, His #status predicted

Query Match 5.1%; Score 6; DB 2; Length 561;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLG 88
 Db 30 KGVVLG 35
 |||||

RESULT 837
 S62788
 carboxylesterase (EC 3.1.1.1) ES-4 precursor, liver - rat
 N;Alternate names: hydrolase B
 C;Species: *Rattus norvegicus* (Norway rat)
 C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C;Accession: S62788; S51203; A55304; S49257
 R;Robbi, M.; van Schaftingen, E.; Beaufay, H.
 Biochem. J. 313, 821-826, 1996
 A;Title: Cloning and sequencing of rat liver carboxylesterase ES-4 (microsomal palmitoyl
 A;Reference number: S62788; MUID:96190723; PMID:8611161
 A;Accession: S62788
 A;Molecule type: mRNA
 A;Residues: 1-561 <ROB>
 A;Cross-references: UNIPROT:Q64573; UNIPARC:UPI000012A1F7; EMBL:X81825; NID:G550417; PID
 A;Experimental source: liver
 R;Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
 Arch. Biochem. Biophys. 315, 495-512, 1994
 A;Title: Purification and characterization of two rat liver microsomal carboxylesterases
 A;Reference number: S51202; MUID:95077430; PMID:7986098
 A;Accession: S51203
 A;Molecule type: protein
 A;Residues: 19-48 <MOR>
 A;Cross-references: UNIPARC:UPI00000E8215
 A;Experimental source: liver
 R;Yan, B.; Yang, D.; Brady, M.; Parkinson, A.
 J. Biol. Chem. 269, 29688-29696, 1994
 A;Title: Rat kidney carboxylesterase. Cloning, sequencing, cellular localization, and re
 A;Reference number: A55304; MUID:95050819; PMID:7961958
 A;Accession: A55304
 A;Molecule type: mRNA
 A;Residues: 1-6, F', 8-59, P', 61-212, A', 214-252, T', 254-309, IT', 312-341, N', 343-424, F'
 A;Cross-references: UNIPARC:UPI00001758F4; GB:U10697; NID:G562007
 A;Note: the sequence in GenBank entry RNUI0697, release 107, (PID:G562008) has the codon
 R;Robbi, M.; Beaufay, H.
 submitted to the EMBL Data Library, September 1994
 A;Reference number: S49257
 A;Accession: S49257
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-561 <RO2>
 A;Cross-references: UNIPARC:UPI000012A1F7; EMBL:X81825; NID:G550417; PIDN:CAA57419.1; P
 C;Function:
 A;Description: catalyzes conversion of carboxylic ester to alcohol and carboxylic anion
 C;Superfamily: cholinesterase; cholinesterase homology
 C;Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; liver
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-561/Product: carboxylesterase ES-4 #status experimental <MAT>
 F;50-551/Domain: cholinesterase homology <CHE>
 F;79,301/Binding site: Ser, His #status predicted
 F;221,466/Active site: Ser, His #status predicted

Query Match 5.1%; Score 6; DB 2; Length 561;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGVVLG 88
|||||
Db 30 KGVVLG 35

RESULT 838

G84244
DNA ligase [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitchauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablid Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84244
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-561 <STO>
A:Cross-references: UNIPROT:Q9HR35; UNIPARC:UPI0000129665; GB:AE004437; NID:gl0580445; E
C:Genetics:
A:Gene: lig
C:Superfamily: DNA ligase

Query Match 5.1%; Score 6; DB 2; Length 561;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 128 LAALAA 133

RESULT 839

T38080
Hypothetical protein SPAC1F3.09 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38080
R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, April 1996
A:Reference number: Z21767
A:Accession: T38080
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-561 <CON>
A:Cross-references: UNIPROT:Q10414; UNIPARC:UPI000013A8CE; EMBL:Z70690; PIDN:CAA94627.1;
A:Experimental source: strain 972h-; cosmid c1F3
C:Genetics:
A:Gene: SPDB.SPAC1F3.09
A:Map position: 1
A:Introns: 12/2; 68/1; 344/1; 371/1

Query Match 5.1%; Score 6; DB 2; Length 561;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 187 LAALAA 192

RESULT 840

JU0033
Hypothetical L1 protein (third intron of gene TS) - human
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 31-Dec-2004
C:Accession: JU0033
R:Horie, N.; Nalbantoglu, J.; Kaneda, S.; Ayusawa, D.; Seno, T.; Takeishi, K.

J. Biochem. 106, 1-4, 1989
A:Title: Identification and characterization of an L1 family sequence with a very long
A:Reference number: JU0033; MUID:89380111; PMID:2476429
A:Accession: JU0033
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-562 <HOR>
A:Cross-references: UNIPROT:O00378; UNIPARC:UPI00001785F5
A:Note: this sequence is similar to human teratocarcinoma L1 RNA species and RNA depend:

Query Match 5.1%; Score 6; DB 2; Length 562;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 OKLEAF 114
|||||
Db 135 OKLEAF 140

RESULT 841

S11175
choline transport protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G3213; protein YGL077c
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S11175; S64084
R:Nikawa, J.I.; Hosaka, K.; Tsukagoshi, Y.; Yamashita, S.
J. Biol. Chem. 265, 15996-16003, 1990
A:Title: Primary structure of the yeast choline transport gene and regulation of its ex
A:Reference number: S11175; MUID:90368823; PMID:2203793
A:Accession: S11175
A:Molecule type: DNA
A:Residues: 1-563 <NIK>
A:Cross-references: UNIPROT:P19807; UNIPARC:UPI000012CA9D; EMBL:J05603; NID:gl71329; PI
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64084
A:Molecule type: DNA
A:Residues: 1-563 <RIE>
A:Cross-references: UNIPARC:UPI000012CA9D; EMBL:Z72599; NID:gl322592; PIDN:CAA96782.1;
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:HNM1; CTRL; MIPS:YGL077c
A:Cross-references: MIPS:YGL077c; SGD:S0003045
A:Map position: 7L
C:Superfamily: choline transport protein
C:Keywords: transmembrane protein
F:91-107/Domain: transmembrane #status predicted <TM1>
F:185-201/Domain: transmembrane #status predicted <TM2>
F:214-230/Domain: transmembrane #status predicted <TM3>
F:257-273/Domain: transmembrane #status predicted <TM4>
F:298-314/Domain: transmembrane #status predicted <TM5>
F:345-361/Domain: transmembrane #status predicted <TM6>
F:401-417/Domain: transmembrane #status predicted <TM7>
F:428-444/Domain: transmembrane #status predicted <TM8>
F:466-482/Domain: transmembrane #status predicted <TM9>

Query Match 5.1%; Score 6; DB 1; Length 563;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CVWIVG 38
|||||
Db 496 CVWIVG 501

RESULT 842

E87631
acyl-CoA dehydrogenase, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: E87631

R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A;Reference number: A87249; MUID:21173698; PMID:11259647
 A;Accession: E87631
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-564 <STO>
 A;Cross-references: UNIPROT:Q9A3W4; UNIPARC:UPI00000C7952; GB:AE005673; NID:G13424741; B
 C;Genetics:
 A;Gene: CC3087

Query Match 5.1%; Score 6; DB 2; Length 564;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 QAOVIA 79
 |||||
 Db 551 QAOVIA 556

RESULT 843

S10367
 carboxylesterase (EC 3.1.1.1) ES-10 precursor, microsomal - rat
 N;Alternate names: hydrolase A
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text_change 09-Jul-2004
 C;Accession: S10367; S12468; S12402; S23460; S14361
 R;Robbi, M.; Beaufay, H.; Octave, J.N.
 Biochem. J. 269, 451-458, 1990
 A>Title: Nucleotide sequence of cDNA coding for rat liver pI 6.1 esterase (ES-10), a carboxylesterase from rat liver
 A;Reference number: S10367; MUID:90351366; PMID:2386485
 A;Accession: S10367
 A;Molecule type: mRNA
 A;Residues: 1-565 <ROB1>
 A;Cross-references: UNIPROT:Q9R135; UNIPARC:UPI00000E7562; EMBL:X51974
 A;Note: 168-Gln, 247-Lys, 423-Met, and 506-Asn were also found
 A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 5 in having 265-Lys
 R;Robbi, M.
 submitted to the EMBL Data Library, February 1990

A;Reference number: S12468
 A;Accession: S12468
 A;Molecule type: mRNA

A;Residues: 1-264,'K',266-565 <ROB2>
 A;Cross-references: UNIPARC:UPI0000047AB1; EMBL:X51974; NID:G56898; PIDN:CAA36236.1; PID
 R;Morgan, E.W.; Yan, B.; Greenway, D.; Petersen, D.R.; Parkinson, A.
 Arch. Biochem. Biophys. 315, 495-512, 1994
 A>Title: Purification and characterization of two rat liver microsomal carboxylesterases
 A;Reference number: S51202; MUID:95077430; PMID:7986098
 A;Accession: S51202

A;Molecule type: protein
 A;Residues: 19-48 <ROB>
 A;Cross-references: UNIPARC:UPI00000E736C
 R;Medda, S.; Proia, R.L.
 Eur. J. Biochem. 206, 801-806, 1992

A>Title: The carboxylesterase family exhibits C-terminal sequence diversity reflecting the
 A;Reference number: S23460; MUID:92299008; PMID:1606562
 A;Accession: S23460

A;Status: preliminary; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-185,'Q',187-422,'M',424-505,'N',507-565 <MED>
 A;Cross-references: UNIPARC:UPI0000170BC6; EMBL:X65296; NID:G57553; PIDN:CAA46391.1; PID
 R;Gaustad, R.; Sletten, K.; Lovhaug, D.; Fonnum, F.
 Biochem. J. 274, 693-697, 1991

A>Title: Purification and characterization of carboxylesterases from rat lung.
 A;Reference number: S14361; MUID:91190080; PMID:2012599
 A;Accession: S14361

A;Molecule type: protein
 A;Residues: 19-26,'D',28-37 <GAU>
 A;Cross-references: UNIPARC:UPI00001758F7
 C;Superfamily: cholinesterase; cholinesterase homology

C;Keywords: carboxylic ester hydrolase; endoplasmic reticulum; glycoprotein; homotrimer
 F;1-18/Domain: signal sequence #status predicted <SIG>
 F;19-565/Product: carboxylesterase #status predicted <MAT>
 F;50-551/Domain: cholinesterase homology <CHS>
 F;79,489/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;221,466/Active site: Ser, His #status predicted

Query Match 5.1%; Score 6; DB 2; Length 565;
 Best Local Similarity 100.0%; Pred. No. 7.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVILG 88
 |||||
 Db 30 KGVILG 35

RESULT 844

TFVFS1
 protein-tyrosine kinase (EC 2.7.1.112) src - avian sarcoma virus S1
 C;Species: avian sarcoma virus S1
 C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 05-Oct-2004
 C;Accession: A25375
 R;Ikawa, S.; Hagino-Yamagishi, K.; Kawai, S.; Yamamoto, T.; Toyoshima, K.
 Mol. Cell. Biol. 6, 2420-2428, 1986
 A>Title: Activation of the cellular src gene by transducing retrovirus.
 A;Reference number: A25375; MUID:87064539; PMID:3097513
 A;Accession: A25375
 A;Molecule type: DNA
 A;Residues: 1-568 <IKA>
 A;Cross-references: UNIPROT:P14084; UNIPARC:UPI00000135F25
 C;Genetics:

A;Gene: src
 C;Superfamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
 C;Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; ph
 F;88-137/Domain: SH3 homology <SH3>
 F;148-245/Domain: SH2 homology <SH2>
 F;265-523/Domain: protein kinase homology <KIN>
 F;273-281/Region: protein kinase ATP-binding motif
 F;2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
 F;295/Active site: Lys #status predicted
 F;416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 568;
 Best Local Similarity 100.0%; Pred. No. 7.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 AYCLSV 31
 |||||
 Db 183 AYCLSV 188

RESULT 845

H82668
 subunit F of alkyl hydroperoxide reductase XF1531 [imported] - *Xylella fastidiosa* (strai
 C;Species: *Xylella fastidiosa*
 C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 31-Dec-2004
 C;Accession: H82668
 R;Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A;Reference number: H82515; MUID:20365717; PMID:10910347
 A;Note: for a complete list of authors see reference number A59328 below
 A;Accession: H82668

A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-571 <SIM>
 A;Cross-references: UNIPROT:Q9PD48; UNIPARC:UPI00000C6085; GB:AE003983; GB:AE003849; NFI
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, I.
 as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Froh

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laigh Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Mencia, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak A.; Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XFI531
A;Superfamily: alkyl hydroperoxide reductase, subunit F; thioredoxin reductase homology

Query Match 5.1%; Score 6; DB 2; Length 571;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
Db 132 VLGLLQ 137

RESULT 846
H83117
probable ATP-binding component of ABC transporter PA4222 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
A;Accession: H83117
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bz adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83117
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-574 <STO>
A;Cross-references: UNIPROT:Q9HWG7; UNIPARC:UPI00000C5C79; GB:AE004839; GB:AE004091; NID:20437337
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4222

Query Match 5.1%; Score 6; DB 2; Length 574;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
Db 385 VLLGGV 390

RESULT 847
B84386
glycine-tRNA synthetase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
A;Accession: B84386
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-576 <STO>
A;Cross-references: UNIPROT:Q9HWM7; UNIPARC:UPI0000063AF9; GB:AE004437; NID:g10581762; E C;Genetics:
A;Gene: glyS
C;Superfamily: glycine-tRNA ligase

Query Match 5.1%; Score 6; DB 2; Length 576;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 383 LAALAA 388

RESULT 848

C82082
penicillin-binding protein 3 VC2407 [imported] - Vibrio cholerae (strain N16961 serogro C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
A;Accession: C82082
R;Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J. chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bacs, S.; Qin, H.; Dragoi, I.; Sellers, I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: C82082
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-580 <HEI>
A;Cross-references: UNIPROT:Q9KPG1; UNIPARC:UPI00000C32B9; GB:AE004310; GB:AE003852; NID:20406833
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC2407
A;Map position: 1
C;Superfamily: penicillin-binding protein 3

Query Match 5.1%; Score 6; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
Db 320 VLAALA 325

RESULT 849

T06380
beta-fructofuranosidase (EC 3.2.1.26) - garden pea
N;Alternate names: cell wall invertase
C;Species: Pisum sativum (garden pea)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
A;Accession: T06380
R;Zhang, L.; Cohn, N.S.; Mitchell, J.P.
submitted to the EMBL Data Library, May 1998
A;Reference number: Z15639
A;Accession: T06380
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-582 <ZHA>
A;Cross-references: UNIPROT:Q43079; UNIPARC:UPI00000AC916; EMBL:AF063246; NID:g3152879; A;Experimental source: cv. Little Marvel
C;Genetics:
A;Gene: bfruct1
C;Function:
C;Superfamily: hydrolyzes sucrose to glucose and fructose
C;Keywords: glycosidase; hydrolase

Query Match 5.1%; Score 6; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GCKPAI 48
Db 129 GCKPAI 134

RESULT 850

S67571
hypothetical protein YDL038c - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein D2726
C/Species: Saccharomyces cerevisiae
C/Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C/Accession: S67571
R/Paulin, L.; Saren, A.M.; Laamanen, P.
submitted to the Protein Sequence Database, July 1996
A/Reference number: S67560
A/Accession: S67571
A/Molecule type: DNA
A/Residues: 1-563 <PAU>
A/Cross-references: UNIPROT:Q99175; UNIPARC:UPI000069D82; EMBL:Z74087; NID:gl1431021; P1
A/Experimental source: strain S288C
C/Genetics:
A/Gene: MIPS:YDL038c
A/Cross-references: SGD:S0002196
A/Map position: 4L
C/Suprafamily: pig submaxillary mucin

Query Match 5.1%; Score 6; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 7.8e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 8 EVTTST 13

Db 9 EVTTST 14

RESULT 851

C70330
conserved hypothetical protein aq_340 - Aquifex aeolicus
C/Species: Aquifex aeolicus
C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C/Accession: C70330
R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov
V.
Nature 392, 353-358, 1998
A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A/Reference number: A70300; MUID:98196666; PMID:9537320
A/Accession: C70330
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-585 <AQF>
A/Cross-references: UNIPROT:O66671; UNIPARC:UPI00005631A; GB:AE000684; NID:g2983009; P1
A/Experimental source: strain VP5
C/Genetics:
A/Gene: aq_340

Query Match 5.1%; Score 6; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 7.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 52 KEVLQY 57

Db 337 KEVLQY 342

RESULT 852

TVFVPR
protein-tyrosine kinase (BC 2.7.1.112) src - avian sarcoma virus PR2257
C/Species: avian sarcoma virus PR2257
C/Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 05-Oct-2004
C/Accession: A30174
R/Geryk, J.; Desele, P.; Barnier, J.V.; Svoboda, J.; Nehyba, J.; Karakoz, I.; Rynditch,
J. Virol. 63, 481-492, 1989
A/Title: Transduction of the cellular src gene and 3' adjacent sequences in avian sarcom
A/Reference number: A30174; MUID:89094972; PMID:2463376
A/Accession: A30174
A/Molecule type: DNA
A/Residues: 1-587 <GER>

A/Cross-references: UNIPROT:P15054; UNIPARC:UPI0000135F23; GB:M21526; NID:g210264; PIDN:
C/Genetics:

A/Gene: src
C/Suprafamily: Tyrosine-protein kinase, proto-oncogene SRC type; protein kinase homology
C/Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; onc
P/88-137/Domain: SH3 homology <SH3>
F/148-245/Domain: SH2 homology <SH2>
F/265-523/Domain: protein kinase homology <KIN>
F/273-281/Region: protein kinase ATP-binding motif
F/2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted
F/295/Active site: Lys #status predicted
F/416/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 5.1%; Score 6; DB 1; Length 587;
Best Local Similarity 100.0%; Pred. No. 7.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 26 AYCLSV 31

Db 183 AYCLSV 188

RESULT 853

B87430
flagellin modification protein FlmG [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C/Accession: B87430
R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: B87430
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-596 <STO>
A/Cross-references: UNIPROT:P21296; UNIPARC:UPI000012A8D8; GB:AE005673; NID:gl3422826; P
C/Genetics:
A/Gene: CC1457

Query Match 5.1%; Score 6; DB 2; Length 596;
Best Local Similarity 100.0%; Pred. No. 8e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

QY 68 AAPYIE 73

Db 385 AAPYIE 390

RESULT 854

S33578
rop protein - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S33578; S33577; S55021
R/Salzberg, A.; Cohen, N.; Halachmi, N.; Kimchie, Z.; Lev, Z.
Development 117, 1309-1319, 1993
A/Title: The Drosophila Ras2 and Rop gene pair: a dual homology with a yeast Ras-like gc
A/Reference number: S33577; MUID:94008534; PMID:8404533
A/Accession: S33578

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-597 <SAL>

A/Cross-references: UNIPROT:Q07327; UNIPARC:UPI000016BD34; EMBL:X67219; NID:g311371; P1
A/Accession: S33577
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-597 <SA2>
A/Cross-references: UNIPARC:UPI000016BD34; EMBL:X67218; NID:g311369; PIDN:CAA47658.1; P1
R/Harrison, S.D.; Solomon, N.; Rubin, G.M.
Genetics 139, 1701-1709, 1995

A:Title: A genetic analysis of the 63E-64A genomic region of *Drosophila melanogaster*: id
A:Reference number: S55020; MUID:95309683; PMID:7789770
A:Accession: S55021
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-523,'S',525-597 <HAR>
A:Cross-references: UNIPARC:UPI00001246ND; EMBL:U15967; NID:G639707; PIDN:AB60242.1; PI
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
C:Genetics:
A:Gene: FlyBase:Rop
A:Cross-references: FlyBase:FBgn0004574
C:Superfamily: vacuolar protein sorting protein VPS45

Query Match 5.1%; Score 6; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVL 55
|||||
Db 284 PDKEVL 289

RESULT 855
C83129
hypothetical protein PA4140 [imported] - *Pseudomonas aeruginosa* (strain PA01)
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
A:Accession: C83129
R:Stover, C.K.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: C83129
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-597 <STO>
A:Cross-references: UNIPROT:Q9HWL1; UNIPARC:UPI00000C5C31; GB:AE004830; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4140

Query Match 5.1%; Score 6; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 193 LGGVLA 198

RESULT 856
S6669
potassium channel (Kv1.5) - rabbit
C:Species: *Oryctolagus cuniculus* (domestic rabbit)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
A:Accession: S6669
R:Sasaki, Y.; Ishii, K.; Nunoki, K.; Yamagishi, T.; Taira, N.
FEBS Lett. 372, 20-24, 1995
A:Title: The voltage-dependent K(+) channel (Kv1.5) cloned from rabbit heart and facilit
A:Reference number: S6669; MUID:96032538; PMID:7556635
A:Accession: S6669
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-598 <SAS>
A:Cross-references: UNIPROT:P50638; UNIPARC:UPI00001279AB; EMBL:D45025; NID:g1060972; PI
C:Superfamily: potassium channel protein drk1

Query Match 5.1%; Score 6; DB 2; Length 598;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 CSQAAP 70
|||||
Db 26 CSQAAP 31

RESULT 857
S23991
conjugative transfer protein traN precursor - *Escherichia coli* plasmid F
C:Species: *Escherichia coli*
C>Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 09-Jul-2004
A:Accession: S23991; B39442; S17095
R:Maneewannakul, S.; Kathir, P.; Ippen-Ihler, K.
J. Mol. Biol. 225, 299-311, 1992
A:Title: Characterization of the F plasmid mating aggregation gene traN and of a new F
A:Reference number: S23990; MUID:92277643; PMID:1593622
A:Accession: S23991
A:Molecule type: DNA
A:Residues: 1-602 <MAN>
A:Cross-references: UNIPROT:P24082; UNIPARC:UPI00001372F4; EMBL:X61575; NID:G43125; PIDN:
R:Maneewannakul, S.; Maneewannakul, K.; Ippen-Ihler, K.
J. Bacteriol. 173, 3872-3878, 1991
A:Title: Characterization of trbC, a new F plasmid tra operon gene that is essential to
A:Reference number: A39442; MUID:91267954; PMID:2050638
A:Accession: B39442
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-19 <MA2>
A:Cross-references: UNIPARC:UPI000017AA35; GB:M60427
C:Genetics:
A:Gene: traN
A:Genome: plasmid
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-602/Product: conjugative transfer protein traN #status predicted <MAT>

Query Match 5.1%; Score 6; DB 2; Length 602;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 425 LAALAA 430

RESULT 858
S01066
regulatory protein nifA - *Bradyrhizobium japonicum*
C:Species: *Bradyrhizobium japonicum*
C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 05-Oct-2004
A:Accession: S01066
R:Thoeny, B.; Fischer, H.M.; Anthamatten, D.; Bruderer, T.; Hennecke, H.
Nucleic Acids Res. 15, 8479-8499, 1987
A:Title: The symbiotic nitrogen fixation regulatory operon (fixRnifA) of *Bradyrhizobium*
A:Reference number: S01065; MUID:88040468; PMID:3313281
A:Accession: S01066
A:Molecule type: DNA
A:Residues: 1-605 <THO>
A:Cross-references: UNIPARC:UPI000016E6FE; EMBL:X06167; NID:G39526; PIDN:CAA29531.1; PI
C:Genetics:
A:Gene: nifA
C:Superfamily: response regulator (sigma54-dependent transcriptional activator), PhlA ty
F:253-474/Domain: RNA polymerase sigma factor interaction domain homology <SFI>

Query Match 5.1%; Score 6; DB 2; Length 605;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||||
Db 77 VLGLLQ 82

RESULT 859
B95322

NodM Glutamine aminotransferase [imported] - Sinorhizobium meliloti (strain 1021) megap1
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C:Accession: B95322
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
L.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11491432
A:Accession: B95322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-608 <KUR>
A:Cross-references: UNIPROT:Q922K3; UNIPARC:UPI0000164826; GB:AE006469; PION:AAK65140.1
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: nodM
A:Genome: plasmid
C:Superfamily: glutamine-fructose-6-phosphate aminotransferase (isomerizing)

Query Match 5.1%; Score 6; DB 2; Length 608;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
|||||
Db 408 VLAALA 413

RESULT 860
A82358
vitamin B12 receptor VC0156 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82358
R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: A82358
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-611 <HEI>
A:Cross-references: UNIPROT:Q9KVI9; UNIPARC:UPI00000C2BF9; GB:AE004105; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0156
A:Map position: 1
C:Superfamily: vitamin B12 receptor; tonB-dependent receptor amino-terminal homology; to

Query Match 5.1%; Score 6; DB 2; Length 611;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 55 LYQQYD 60
|||||
Db 414 LYQQYD 419

RESULT 861
T03890
hypothetical protein C13D9.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 31-Dec-2004
C:Accession: T03890
R:Nelson, J.; Wohldmann, P.; Beck, C.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid C13D9.
A:Reference number: Z15128
A:Accession: T03890
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-611 <NEL>
A:Cross-references: UNIPROT:O16243; UNIPARC:UPI00000748F1; EMBL:AF016420; NID:g22291168;
C:Genetics:
A:Map position: V
A:Introns: 70/3; 124/1; 180/1; 268/3; 346/2; 394/1; 449/1
A:Note: C13D9.9

Query Match 5.1%; Score 6; DB 2; Length 611;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIV 49
|||||
Db 472 GKPAIV 477

RESULT 862
C90121
DNA repair helicase [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: C90121
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Re
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: C90121
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-617 <DOU>
A:Cross-references: UNIPROT:Q98SB9; UNIPARC:UPI000008DD05; GB:AF083031; NID:g13794287; F
C:Genetics:
A:Gene: rad25
A:Map position: 3
A:Genome: nucleomorph
C:Keywords: nucleomorph

Query Match 5.1%; Score 6; DB 2; Length 617;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 85 KVLGLL 90
|||||
Db 353 KVLGLL 358

RESULT 863
S38923
hypothetical protein 11 - phage phi-C31
C:Species: phage phi-C31
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C:Accession: S38923
R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.P.
submitted to the EMBL Data Library, November 1993
A:Reference number: S38912
A:Accession: S38923
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-617 <HAR>
A:Cross-references: UNIPROT:Q38032; UNIPARC:UPI000009AF53; EMBL:X76288; NID:g432610; PID
C:Genetics:
A:Start codon: GTG

Query Match 5.1%; Score 6; DB 2; Length 617;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 411 VLAALA 416

RESULT 864
B71559
probable metalloproteinase - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
C:Accession: B71559
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis
A:Reference number: A71570; MUID:99000809; PMID:9784136
A:Accession: B71559
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <ARN>
A:Cross-references: UNIPROT:O84075; UNIPARC:UPI00001391F3; GB:AE001282; GB:AE001273; NID
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics:
A:Gene: yaeL

Query Match 5.1%; Score 6; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 7 VLAALA 12

RESULT 865
H81712
conserved hypothetical protein TC0344 [imported] - Chlamydia muridarum (strain Nigg)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81712
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: A81500; MUID:20150255; PMID:10684935
A:Accession: H81712
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-619 <RET>
A:Cross-references: UNIPROT:Q9PKW7; UNIPARC:UPI00000578B7; GB:AE002302; GB:AE002160; NID
A:Experimental source: strain Nigg (MoPn)
C:Genetics:
A:Gene: TC0344

Query Match 5.1%; Score 6; DB 2; Length 619;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 7 VLAALA 12

RESULT 866
T33969
hypothetical protein F46E10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33969

R;Johnson, D.; Bradshaw, H.
submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid F46E10.
A:Reference number: Z21446
A:Accession: T33969
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-623 <JOH>
A:Cross-references: UNIPROT:Q9UAV8; UNIPARC:UPI0000078B2A; EMBL:AF125955; PIDN:AAI4712
A:Experimental source: strain Bristol N2; clone F46E10
C:Genetics:
A:Gene: CESP:F46E10.1
A:Map position: 5
A:Introns: 14/3; 30/1; 102/3; 145/3; 185/3; 223/2; 533/3; 572/3
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology
F:114-603/Domain: acetate-CoA ligase homology <ACL>

Query Match 5.1%; Score 6; DB 2; Length 623;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
DB 313 GVLAAL 318

RESULT 867
C64180
hypothetical protein HI1056 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 05-Oct-2004
C:Accession: C64180
R:Pfeilschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Pine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M. Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: C64180
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-629 <TIG>
A:Cross-references: UNIPROT:P71366; UNIPARC:UPI00000512C9; GB:U32786; GB:I42023; NID:gt
C:Superfamily: type III restriction-modification system methyltransferase

Query Match 5.1%; Score 6; DB 2; Length 629;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QPKGV 86
|||||
DB 138 QPKGV 143

RESULT 868
S32349
probable SNP2-type helicase - Bacillus cereus (fragment)
C:Species: Bacillus cereus
C>Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 09-Jul-2004
C:Accession: S32349
R:Kolsto, A.B.; Bork, P.; Kvaloy, K.; Lindback, T.; Gronstadt, A.; Kristensen, T.; Sande J. Mol. Biol. 230, 684-698, 1993
A:Title: Prokaryotic members of a new family of putative helicases with similarity to tr
A:Reference number: S32349; MUID:93218008; PMID:8464078
A:Accession: S32349
A:Molecule type: DNA
A:Residues: 1-634 <KOL>
A:Cross-references: UNIPROT:P94295; UNIPARC:UPI000017AC8E
A:Experimental source: strain ATCC 10987
C:Genetics:
A:Gene: snf
C:Keywords: DNA binding

```
Query Match      5.1%; Score 6; DB 2; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19  GVLAAL 24
      |||||
Db      354  GVLAAL 359

RESULT 869
T35498
probable anthranilate synthase - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T35498
R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21580
A:Accession: T35498
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-634 <SEE>
A:Cross-references: UNIPROT:Q9S2M6; UNIPARC:UPI00000DB2BD; EMBL:AL109661; PIDN:CAB51965.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: trpB1; SCOREDB:SC6E10.11

Query Match      5.1%; Score 6; DB 2; Length 634;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19  GVLAAL 24
      |||||
Db      376  GVLAAL 381

RESULT 870
AC3099
conserved hypothetical protein Atu4415 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AC3099
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.;
et al.; Gilllet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan,
et al.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AC3099
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-637 <KUR>
A:Cross-references: UNIPROT:Q8U7N4; UNIPARC:UPI0000164925; GB:AE008689; PIDN:AAL45209.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu4415
A:Map position: linear chromosome
C:Superfamily: Campylobacter jejuni hypothetical protein Cj0145

Query Match      5.1%; Score 6; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17  LGGVLA 22
      |||||
Db      41  LGGVLA 46

RESULT 871
D71944
```

```
transketolase - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 09-Jul-2004
C:Accession: D71944
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: D71944
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-641 <ARN>
A:Cross-references: UNIPROT:Q9ZW85; UNIPARC:UPI00000D71B6; GB:AE001469; GB:AE001439; NID
A:Experimental source: strain J99
C:Genetics:
A:Gene: tkTA
C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology

Query Match      5.1%; Score 6; DB 2; Length 641;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      82  FKGVKL 87
      |||||
Db      583  FKGVKL 588

RESULT 872
H64655
transketolase A - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 09-Jul-2004
C:Accession: H64655
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: H64655
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-641 <TOM>
A:Cross-references: UNIPROT:O25720; UNIPARC:UPI00000D3198; GB:AE000615; GB:AE000511; NID
C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology

Query Match      5.1%; Score 6; DB 2; Length 641;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      82  FKGVKL 87
      |||||
Db      583  FKGVKL 588

RESULT 873
FOVZZW
major core protein P4b precursor - vaccinia virus (strain WR)
C:Species: vaccinia virus
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 09-Jul-2004
C:Accession: A03871
R:Roel, J.; Moss, B.
J. Virol. 56, 830-838, 1985
A:Title: Transcriptional and translational mapping and nucleotide sequence analysis of a
A:Reference number: A03871; MUID:86062913; PMID:2999438
A:Accession: A03871
A:Molecule type: DNA
A:Residues: 1-643 <ROS>
A:Cross-references: UNIPROT:P06440; UNIPARC:UPI0000138C37; GB:M11079; NID:G335714; PIDN
C:Superfamily: vaccinia virus major core protein P4b
```

C:Keywords: core protein
F:1-61/Domain: leader peptide #status predicted <LDR>
F:62-643/Product: major core protein P4b #status predicted <MAT>

Query Match 5.1%; Score 6; DB 1; Length 643;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QAPYI 72
Db 382 QAPYI 387

RESULT 874

GB7626
DNA primase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: GB7626
R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249, PMID:21173698; PMID:11259647
A:Accession: GB7626
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-643 <STO>
A:Cross-references: UNIPROT:Q9A401; UNIPARC:UPI00000C792D; GB:AE005673; NID:gl3424695; F
A:Gene: CC3049

Query Match 5.1%; Score 6; DB 2; Length 643;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
Db 547 GVLAAL 552

RESULT 875

FOVZSR
major core protein P4b precursor - vaccinia virus (strain Copenhagen)
N:Alternate names: A3L protein
C:Species: vaccinia virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C:Accession: E42517
R.Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A:Title: Appendix to "The complete DNA sequence of vaccinia virus".
A:Reference number: A42501
A:Accession: E42517
A:Molecule type: DNA
A:Residues: 1-644 <GOE>
A:Cross-references: UNIPROT:P20643; UNIPARC:UPI0000138C36; GB:M35027; NID:g335317; PIDN:
A:Experimental source: strain Copenhagen
R.Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 247-266, 1990
A:Title: The complete DNA sequence of vaccinia virus.
A:Reference number: A42531; PMID:91021027; PMID:2219722
A:Contents: annotation; possible protein-coding frames
A:Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: vaccinia virus major core protein P4b
C:Keywords: core protein
F:1-61/Domain: leader peptide #status predicted <LDR>
F:62-644/Product: major core protein P4b #status predicted <MAT>

Query Match 5.1%; Score 6; DB 1; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QAPYI 72
Db 382 QAPYI 387

RESULT 876

T37390
major core protein P4b - vaccinia virus (strain Ankara)
C:Species: vaccinia virus
A:Variety: strain Ankara
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T37390
R:Antoine, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.
submitted to the EMBL Data Library, March 1997
A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) stra
A:Reference number: Z20877
A:Accession: T37390
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-644 <ANT>
A:Cross-references: UNIPROT:O93121; UNIPARC:UPI00000F48B2; EMBL:U94848; PIDN:AAB96457.1
A:Experimental source: strain Ankara
C:Genetics:
A:Note: MVA114L
C:Superfamily: vaccinia virus major core protein P4b

Query Match 5.1%; Score 6; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QAPYI 72
Db 382 QAPYI 387

RESULT 877

A72164
A4L protein - variola minor virus (strain Garcia-1966)
C:Species: variola minor virus
C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
C:Accession: A72164
R.Shchelkunov, S.N.; Totmenin, A.V.; Gutorov, V.V.; Safronov, P.F.; Massung, R.F.; Lopa:
submitted to GenBank, March 1998
A:Description: Analysis of the complete coding sequence of DNA of alastrim variola mino
A:Reference number: A72150
A:Accession: A72164
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-644 <SHC>
A:Cross-references: UNIPROT:P33818; UNIPARC:UPI0000000750; GB:Y16780; NID:g5830555; PIDN:
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: A4L
C:Superfamily: vaccinia virus major core protein P4b

Query Match 5.1%; Score 6; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QAPYI 72
Db 382 QAPYI 387

RESULT 878

D36848
major core protein P4b - variola virus
N:Alternate names: A3L protein
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
A:Accession: D36848; S46890
R:Blinov, V.M.

```

submitted to GenBank, November 1992
A:Reference number: A36859
A:Accession: D36848
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-644 <BLI>
A:Cross-references: UNIPROT:P33818; UNIPARC:UPI0000000750; GB:X69198; NID:g456758; PIDN:
A:Experimental source: strain India-1967
R:Volchov, V.E.; Blinov, V.M.; Totmenin, A.V.; Shchelkunov, S.N.; Sandakhchiev, L.S.
Submitted to the EMBL Data Library, April 1992
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI-G genome
A:Reference number: S46890
A:Accession: S46890
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-175 <VOL>
A:Cross-references: UNIPARC:UPI000000EP23D; EMBL:X67116; NID:g516451; PIDN:CAA47512.1; PI
C:Superfamily: vaccinia virus major core protein P4b

Query Match      5.1%; Score 6; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      67 QAPYI 72
      |||||
Db      382 QAPYI 387

RESULT 879
T28545
hypothetical protein A4L - variola major virus
C:Species: variola major virus
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
A:Accession: T28545
R:Masung, R.F.; Eposito, J.J.; Liu, L.I.; Qi, J.; Utterback, T.R.; Knight, J.C.; Aubin
Nature 366, 748-751, 1993
A>Title: Potential virulence determinants in terminal regions of variola smallpox virus
A:Reference number: Z20488; MUID:94088747; PMID:8264798
A:Accession: T28545
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-644 <MAS>
A:Cross-references: UNIPROT:P33818; UNIPARC:UPI0000000750; EMBL:L22579; NID:g6233595; PID
A:Experimental source: strain Bangladesh-1975
C:Superfamily: vaccinia virus major core protein P4b

Query Match      5.1%; Score 6; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      67 QAPYI 72
      |||||
Db      382 QAPYI 387

RESULT 880
G82490
methyl-accepting chemotaxis protein VCA0176 [imported] - Vibrio cholerae (strain N16961
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
A:Accession: G82490
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: G82490
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-644 <HEI>
A:Cross-references: UNIPROT:Q9KMZ0; UNIPARC:UPI000000C345E; GB:AE004358; GB:AE003853; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor

```

```

C:Genetics:
A:Gene: VCA0176
A:Map position: 2

Query Match      5.1%; Score 6; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      8 EVTST 13
      |||||
Db      612 EVTST 617

RESULT 881
S19156
serotonin receptor 2B - fruit fly (Drosophila melanogaster)
N:Alternate names: 5-hydroxytryptamine receptor 2B (5-HTR2B)
C:Species: Drosophila melanogaster
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
A:Accession: S19156; S18154
R:Saudou, F.; Boschart, U.; Amlaiky, N.; Plassat, J.L.; Hen, R.
EMBO J. 11, 7-17, 1992
A>Title: A family of Drosophila serotonin receptors with distinct intracellular signalling
A:Reference number: S19155; MUID:92155185; PMID:1310937
A:Accession: S19156
A:Molecule type: mRNA
A:Residues: 1-645 <SAU>
A:Cross-references: UNIPROT:P28286; UNIPARC:UPI0000124F41; EMBL:Z11490; NID:g7506; PIDN:
C:Genetics:
A:Gene: FlyBase:5-HT1B
A:Cross-references: FlyBase:PBgn0004572
C:Superfamily: octopamine receptor type I
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:124-145/Domain: transmembrane #status predicted <TM1>
F:156-177/Domain: transmembrane #status predicted <TM2>
F:193-214/Domain: transmembrane #status predicted <TM3>
F:234-256/Domain: transmembrane #status predicted <TM4>
F:284-305/Domain: transmembrane #status predicted <TM5>
F:564-587/Domain: transmembrane #status predicted <TM6>
F:597-619/Domain: transmembrane #status predicted <TM7>

Query Match      5.1%; Score 6; DB 2; Length 645;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 22
      |||||
Db      531 LGGVLA 536

RESULT 882
T42296
hypothetical protein - phage SPPI
C:Species: phage SPPI
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
A:Accession: T42296
R:Alonso, J.C.; Luder, G.; Stiege, A.C.; Chai, S.; Weise, F.; Trautner, T.A.
Gene 204, 201-212, 1997
A>Title: The complete nucleotide sequence and functional organization of Bacillus subtil
A:Reference number: Z22137; MUID:98094274; PMID:9434185
A:Accession: T42296
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-646 <ALO>
A:Cross-references: UNIPROT:O48456; UNIPARC:UPI000009B4CA; EMBL:X97918; PIDN:CAA66557.1

Query Match      5.1%; Score 6; DB 2; Length 646;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 LAALAA 26
      |||||
Db      242 LAALAA 247

```

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: C70688
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-652 <COL>
 A:Cross-references: UNIPROT:P71994; UNIPARC:UPI00000D6020; GB:Z81360; GB:AL123456; NID:
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: narX

Query Match 5.1%; Score 6; DB 2; Length 652;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VVIVGH 39
 |||||
 DB 474 VVIVGH 479

RESULT 886
 AH2445
 hypothetical protein all5120 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp. PCC 7120
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C:Accession: AH2445
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AH2445
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-654 <KUR>
 A:Cross-references: UNIPROT:Q8YM21; UNIPARC:UPI00000CED7B; GB:BA000019; PIDN:BA076819.1;
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all5120

Query Match 5.1%; Score 6; DB 2; Length 654;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
 |||||
 DB 97 VLLGGV 102

RESULT 887
 S43415
 histidine ammonia-lyase (EC 4.3.1.3) - human
 N:Alternate names: histidase
 C:Species: Homo sapiens (man)
 C:Date: 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
 C:Accession: S43415
 R:Suchi, M.; Harada, N.; Wada, Y.; Takagi, Y.
 Biochim. Biophys. Acta 1216, 293-295, 1993
 A:Title: Molecular cloning of a cDNA encoding human histidase.
 A:Reference number: S43415; MUID:94060103; PMID:7916645
 A:Accession: S43415
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-657 <SUC>
 A:Cross-references: UNIPROT:P42357; UNIPARC:UPI000012CE8E; DDBJ:U16626; NID:g451209; PII
 C:Genetics:
 A:Gene: GDB:HAL; HIS
 A:Cross-references: GDB:120746; OMIM:235800
 A:Map position: 12q22-12q23
 C:Function:
 A:Description: catalyzes the formation of (E)-3-(1H-imidazol-4-yl)-propenoic acid (uroc
 A:Pathway: histidine catabolism

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: G72346
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-647 <ARN>
 A:Cross-references: UNIPROT:Q9WZ80; UNIPARC:UPI00000C13B5; GB:AE001740; GB:AE000512; NID
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM0672

Query Match 5.1%; Score 6; DB 2; Length 647;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGKVLG 88
 |||||
 DB 44 KGKVLG 49

RESULT 884
 AG0314
 probable thiamin pyrophosphate-dependent protein YPO2578 [imported] - Yersinia pestis (s
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 31-Dec-2004
 C:Accession: AG0314
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: AG0314
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-648 <KUR>
 A:Cross-references: UNIPROT:Q8ZDI8; UNIPARC:UPI00000CD93B; GB:AL590842; PIDN:CAC91379.1;
 C:Genetics:
 A:Gene: YPO2578
 C:Superfamily: thiamine diphosphate-dependent enzyme, acetolactate synthase type; thiami

Query Match 5.1%; Score 6; DB 2; Length 648;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 DB 586 LAALAA 591

RESULT 885
 C70688
 probable nitrate reductase - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
 C:Accession: C70688
 R:Colé, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

C;Superfamily: histidine ammonia-lyase
C;Keywords: ammonia-lyase; carbon-nitrogen lyase; histidine catabolism
F;253/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted
F;254/Modified site: dehydroalanine (Ser) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 8.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 68 AAPYIE 73
| | | | |
Db 616 AAPYIE 621

RESULT 888
A46128
histidine ammonia-lyase (EC 4.3.1.3) - mouse
N;Alternate names: histidase
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A46128
R;Taylor, R.G.; Grieco, D.; Clarke, G.A.; McInnes, R.R.; Taylor, B.A.
Genomics 16, 231-240, 1993
A;Title: Identification of the mutation in murine histidinemia (his) and genetic mapping
A;Reference number: A46128; MUID:93252384; PMID:8486363
A;Accession: A46128
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-657 <TAV>
A;Cross-references: UNIPROT:P35492; UNIPARC:UPI0000003F1B; GB:I07645; NID:g193751; PIDN:
A;Note: sequence extracted from NCBI backbone (NCBIN:131641, NCBIP:131644)
C;Function:
A;Description: catalyzes the formation of (E)-3-(1H-imidazol-4-yl)-propenoic acid (uroc
A;Pathway: histidine catabolism
C;Superfamily: histidine ammonia-lyase
C;Keywords: ammonia-lyase; carbon-nitrogen lyase; histidine catabolism
F;253/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted
F;254/Modified site: dehydroalanine (Ser) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 8.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 68 AAPYIE 73
| | | | |
Db 616 AAPYIE 621

RESULT 889
A36087
histidine ammonia-lyase (EC 4.3.1.3) - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 09-Jul-2004
C;Accession: A36087
R;Taylor, R.G.; Lambert, M.A.; Sexsmith, E.; Sadler, S.J.; Ray, P.N.; Mahuran, D.J.; McI
J. Biol. Chem. 265, 18192-18199, 1990
A;Title: Cloning and expression of rat histidase. Homology to two bacterial histidases a
A;Reference number: A36087; MUID:91009306; PMID:2120224
A;Accession: A36087
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-657 <TAV>
A;Cross-references: UNIPROT:P21213; UNIPARC:UPI00000471C3; GB:M58308; NID:g20
C;Superfamily: histidine ammonia-lyase
C;Keywords: ammonia-lyase; carbon-nitrogen lyase; histidine catabolism
F;253/Cross-link: 5-imidazolinone (Ala-Gly) #status predicted
F;254/Modified site: dehydroalanine (Ser) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 8.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 68 AAPYIE 73
| | | | |
Db 616 AAPYIE 621

RESULT 890
T46359
hypothetical protein DKFZp434Kl316.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C;Accession: T46359
R;Koeherer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23037
A;Accession: T46359
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-658 <AAA>
A;Cross-references: UNIPROT:Q9WTF9; UNIPARC:UPI0000007139F; EMBL:AL137291
A;Experimental source: adult testis; clone DKFZp434Kl316
C;Genetics:
A;Note: DKFZp434Kl316.1

Query Match 5.1%; Score 6; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 8.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 20 VLAALA 25
| | | | |
Db 62 VLAALA 67

RESULT 891
D84286
3-hydroxyacyl-CoA dehydrogenase [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84286
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84286
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-659 <STO>
A;Cross-references: UNIPROT:Q9HQ61; UNIPARC:UPI0000063864; GB:AE004437; NID:g10580830; P
C;Genetics:
A;Gene: hbd2
C;Superfamily: probable 3-hydroxyacyl-CoA dehydrogenase; 3-hydroxyacyl-CoA dehydrogenase

Query Match 5.1%; Score 6; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 8.7e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;

Qy 37 VGHIEL 42
| | | | |
Db 413 VGHIEL 418

RESULT 892
F98187
hypothetical protein AGR_L_902 [imported] - Agrobacterium tumefaciens (strain C58, Cerec
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: F98187
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F98187
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-659 <RUR>
C;Cross-references: UNIPROT:Q8U7N4; UNIPARC:UPI00000D2570; GB:AE007870; PIDN:AAK89024.1;
C;Genetics:
A;Gene: AGR_L_902
A;Map position: linear chromosome
C;Superfamily: Campylobacter jejuni hypothetical protein Cj0145

Query Match 5.1%; Score 6; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||
Db 63 LGGVLA 68

RESULT 893
F83342
probable cation-transporting P-type ATPase PA2435 [imported] - Pseudomonas aeruginosa (e
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: F83342
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: F83342
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-661 <STO>
A;Cross-references: UNIPROT:Q9I147; UNIPARC:UPI00000C56CF; GB:AE004670; GB:AE004091; NID
F84377
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2435
C;Superfamily: Enterococcus copper-transporting ATPase copB; ATPase nucleotide-binding o

Query Match 5.1%; Score 6; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||
Db 388 LAALAA 393

RESULT 894
S42799
garp precursor - human
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: S42799; I37407
R;Birbaum, D.
submitted to the EMBL Data Library, July 1993
A;Reference number: S42799
A;Accession: S42799
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-662 <BIR>
A;Cross-references: UNIPROT:Q14392; UNIPARC:UPI000012B0DF
R;Ollendorff, V.; Noguchi, T.; Delapeyriere, O.; Birbaum, D.
Cell Growth Differ. 5, 213-219, 1994
A;Title: The GARP gene encodes a new member of the family of leucine-rich repeat-contain
A;Reference number: I37407; MUID:94235567; PMID:8180135
A;Accession: I37407
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-662 <RES>

A;Cross-references: UNIPARC:UPI000012B0DF; EMBL:Z24680; NID:g439295; PIDN:CAA80847.1; P
C;Genetics:
A;Gene: GDB:GARP; D11S833E
A;Cross-references: GDB:433911
A;Map position: 11q13.5-11q14
P;50-73/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;74-97/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;98-121/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;150-173/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;174-197/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;198-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;219-242/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;244-265/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F;266-289/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>
F;316-339/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR11>
F;340-363/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR12>
F;364-386/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR13>
F;387-410/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR14>
F;411-433/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR15>
F;444-466/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR16>
F;467-490/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR17>
F;492-514/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR18>
F;515-536/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR19>
F;537-560/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR20>

Query Match 5.1%; Score 6; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||
Db 37 VLGLLQ 42

RESULT 895
F84377
mismatch repair protein [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 12-Jul-2004
C;Accession: F84377
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithausner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; L
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: F84377
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-669 <STO>
A;Cross-references: UNIPROT:Q9HN35; UNIPARC:UPI0000063AC1; GB:AE004437; NID:g10581681;
C;Genetics:
A;Gene: mutS3
C;Superfamily: DNA mismatch repair protein, archaeal type

Query Match 5.1%; Score 6; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||
Db 77 VLGLLQ 82

RESULT 896
A87441
penicillin-binding protein 2 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87441
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
A;Accession: AB7441
A;Reference number: AB7249; MUID:21173698; PMID:11259647

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-672 <STO>
A;Cross-references: UNIPROT:Q9A818; UNIPARC:UPI00000C7410; GB:AE005673; NID:gl3422931; F
C;Genetics:
A;Gene: CC1546
C;Superfamily: penicillin-binding protein 3

Query Match 5.1%; Score 6; DB 2; Length 672;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 344 LAALAA 349

RESULT 897

GB7586
cation transporting ATPase CC2726 [imported] - *Caulobacter crescentus*
C;Species: *Caulobacter crescentus*
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: GB7586
R;Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of *Caulobacter crescentus*.
A;Reference number: AB7249; MUID:21173698; PMID:11259647

A;Accession: GB7586
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-677 <STO>
A;Cross-references: UNIPROT:Q9A4U9; UNIPARC:UPI00000C780D; GB:AE005673; NID:gl3424315; F
C;Genetics:
A;Gene: CC2726
C;Superfamily: probable cadmium-transporting ATPase F6G3.150; ATPase nucleotide-binding

Query Match 5.1%; Score 6; DB 2; Length 677;
Best Local Similarity 100.0%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 35 LGGVLA 40

RESULT 898

D94248
hypothetical protein Vng0920h [imported] - *Halobacterium* sp. NRC-1
C;Species: *Halobacterium* sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84248
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A>Title: Genome sequence of *Halobacterium* species NRC-1.
A;Reference number: AB4160; MUID:20504483; PMID:11016950

A;Accession: D84248
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-688 <STO>
A;Cross-references: UNIPROT:Q9HR07; UNIPARC:UPI0000063773; GB:AE004437; NID:gl0580480; F
C;Genetics:
A;Gene: VNG0920H

Query Match 5.1%; Score 6; DB 2; Length 688;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
DB 304 LGGVLA 309

RESULT 899

T29772
hypothetical protein ZC581.9 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T29772
R;Waterston, B.; Gattung, S.; Le, T.T.
submitted to the EMBL Data Library, May 1997
A;Description: The sequence of *C. elegans* cosmid ZC581.
A;Reference number: Z20682

A;Accession: T29772
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-689 <WAT>
A;Cross-references: UNIPROT:O01776; UNIPARC:UPI000017BCE5; EMBL:AF003134; PIDN:AA054146.
A;Experimental source: strain Bristol N2; clone ZC581
C;Genetics:
A;Gene: CESP:ZC581.9
A;Map position: 1
A;Introns: 43/3; 89/3; 143/1; 222/3; 306/3; 347/3; 430/1; 633/2

Query Match 5.1%; Score 6; DB 2; Length 689;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 392 LAALAA 397

RESULT 900

H75469
conserved hypothetical protein - *Deinococcus radiodurans* (strain R1)
C;Species: *Deinococcus radiodurans*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75469
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A>Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: H75469
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-690 <WHI>
A;Cross-references: UNIPROT:Q9RW25; UNIPARC:UPI000000C1823; GB:AE001938; NID:AE000513; NID
A;Experimental source: strain R1
C;Genetics:
A;Map position: 1
A;Gene: DR0844
C;Superfamily: conserved hypothetical protein yyaL

Query Match 5.1%; Score 6; DB 2; Length 690;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAAAL 25
|||||
DB 436 VLAAAL 441

RESULT 901

G82123
Flagellar biosynthetic protein FlhA VC2069 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 05-Oct-2004
C:Accession: G82123
R:Heidelberg, J.P.; Eiken, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardon, D.; Emolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E.
I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: G82123
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-697 <HEI>
A:Cross-references: UNIPROT:Q9KQD1; UNIPARC:UPI00000C31C9; GB:AE004280; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2069
A:Map position: 1
C:Superfamily: Type III secretion system/flagellar apparatus protein, Inva/LcrD/FlhA type
Query Match 5.1%; Score 6; DB 2; Length 697;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALA 25
|||||
Db 28 VLAALA 33
RESULT 902
WPBSGS
phosphotransferase system enzyme II (EC 2.7.1.69), glucose-specific, factor II [validate]
N:Alternate names: glucose permease; phosphoenolpyruvate:glucose phosphotransferase system
II; PTS glucose-specific enzyme IIBC component ptsG
C:Species: Bacillus subtilis
C>Date: 31-Mar-1991 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: S25083; A36101; S15272; S04174; S04175; A57142; E69683; S22752
R:Zagorec, M.; Postma, P.W.
Mol. Gen. Genet. 234, 325-328, 1992
A:Title: Cloning and nucleotide sequence of the ptsG gene of *Bacillus subtilis*.
A:Reference number: S25083; MUID:92375001; PMID:1508157
A:Accession: S25083
A:Molecule type: DNA
A:Residues: 1-699 <ZAG>
A:Cross-references: UNIPROT:P20166; UNIPARC:UPI00000602FB; EMBL:Z11744; NID:G39955; PIDN
R:Butlira, S.L.; Reddy, P.; Saier Jr., M.H.; Reizer, J.
J. Biol. Chem. 265, 18581-18589, 1990
A:Title: The glucose permease of *Bacillus subtilis* is a single polypeptide chain that fu
A:Reference number: A36101; MUID:91009360; PMID:2120236
A:Accession: A36101
A:Molecule type: DNA
A:Residues: 483-558 <SUT>
A:Cross-references: UNIPARC:UPI000016E892; GB:M60344; NID:G143017; PIDN:AAA22498.1; PID:
R:Gony-Trebol, G.; de Waard, J.H.; Zagorec, M.; Postma, P.W.
Mol. Microbiol. 5, 1241-1249, 1991
A:Title: The glucose permease of the phosphotransferase system of *Bacillus subtilis*: evi
A:Reference number: S15272; MUID:92065821; PMID:1956301
A:Accession: S15272
A:Molecule type: DNA
A:Residues: 249-699 <MOL>
A:Cross-references: UNIPARC:UPI0000172641; EMBL:Z11744; NID:G39955
R:Gony-Trebol, G.; Zagorec, M.; Rain-Guion, M.C.; Steinmetz, M.
Mol. Microbiol. 3, 103-112, 1989
A:Title: Phosphoenolpyruvate:sugar phosphotransferase system of *Bacillus subtilis*: nucle
A:Reference number: S04174; MUID:89237891; PMID:2497294
A:Accession: S04174
A:Molecule type: DNA
A:Residues: 361-517, 'SLDSLRKHLK' <GOL>
A:Cross-references: UNIPARC:UPI000016E921; EMBL:X12832; NID:G48679; PIDN:CAA31315.1; PID
A:Accession: S04175
A:Molecule type: DNA

A:Residues: 'M',539-699 <GO2>
A:Cross-references: UNIPARC:UPI00000B881A; EMBL:X12832; NID:G48679; PIDN:CAA31316.1; PID
A:Note: these sequences have been revised in reference S15272
R:Tolner, B.; Ubink-Kok, T.; Poolman, B.; Konings, W.N.
J. Bacteriol. 177, 2863-2869, 1995
A:Title: Characterization of the proton/glutamate symport protein of *Bacillus subtilis*:
A:Reference number: A57142; MUID:95270606; PMID:7751298
A:Accession: A57142
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 568-572, 'E',574-577, 'E',579-582, 'I',584-587, 'K',589-590, 'A',592, 'AD',595, 'E'
'KQ',646, 'EL',649-650, 'TF',653-654, 'NV',658, 'QRAA',663, 'AI',666-667, 'VI',670-672, 'TS',
A:Cross-references: UNIPARC:UPI000016E765; GB:U15147; NID:9558838; PIDN:AAA82877.1; PID
R:Liao, D.I.; Herzberg, O.
submitted to the Brookhaven Protein Data Bank, September 1991
A:Reference number: A51205; PDB:1GPR
A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 541-698
R:Kapadia, G.; Chen, C.C.H.; Reddy, P.; Saier Jr., M.H.; Reizer, J.; Herzberg, O.
J. Mol. Biol. 221, 1079-1080, 1991
A:Title: Crystallization of the IIA domain of the glucose permease of *Bacillus subtilis*
A:Reference number: A58685; MUID:92046050; PMID:1942043
A:Contents: annotation
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cha
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.;
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, I
A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69683
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-699 <KUN>
A:Cross-references: UNIPARC:UPI00000602FB; GB:299111; GB:AL009126; NID:G2633699; PIDN: C
A:Experimental source: strain 168
C:Genetics:
A:Gene: ptsG
C:Function:
A:Description: transfers a phosphate group from phosphocarryer protein to a sugar which
C:Superfamily: phosphotransferase system N-acetylglucosamine-specific enzyme II; phosph
nzyme II, factor III homology
C:Keywords: phosphohistidine; phosphoprotein; phosphotransferase; sugar transport syste
F:1-517/Domain: phosphotransferase system glucose-specific enzyme II, factor II homology
F:546-699/Domain: phosphotransferase system glucose-specific enzyme II, factor III homology
F:461/Active site: Cys (phosphocysteine intermediate) #status predicted
F:620/Active site: His (phosphohistidine intermediate) #status predicted
Query Match 5.1%; Score 6; DB 1; Length 699;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GVLAAL 24
|||||
Db 153 GVLAAL 158
RESULT 903
T20892
hypothetical protein F14E5.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C:Accession: T20892
R:Percy, C.
submitted to the EMBL Data Library, October 1995

A;Reference number: Z19341
A;Accession: T20892
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-701 <WIL>
A;Cross-references: UNIPROT:Q19460; UNIPARC:UPI000017B8P4; EMBL:Z66522; PIDN:CAA91406.1;
A;Experimental source: clone F14E5
C;Genetics:
A;Gene: CESP:F14E5.3
A;Map position: 2
A;Introns: 39/3; 84/3; 113/3; 186/3; 217/3; 296/3; 384/2; 464/3; 491/3; 531/3; 544/3; 59
C;Superfamily: mammalian acid phosphatase

Query Match 5.1%; Score 6; DB 2; Length 701;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
Db 221 LLGGVL 226
|||||

RESULT 904
S25576
probable transport protein mt2 - rat
N;Alternate names: TAP2 protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 05-Oct-2004
A;Accession: S25576; S21740; S38401
R;Povis, S.J.; Deverson, E.V.; Coadwell, W.J.; Ciruela, A.; Huskisson, N.S.; Smith, H.;
Nature 357, 211-215, 1992
A;Title: Effect of polymorphism of an MHC-linked transporter on the peptides assembled i
A;Reference number: S21740; MUID:92269933; PMID:1350326
A;Accession: S25576
A;Molecule type: mRNA
A;Residues: 1-703 <POW>
A;Cross-references: UNIPROT:Q63521; UNIPARC:UPI000008890E
A;Accession: S21740
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-45,'L',47-703 <POJ>
A;Cross-references: UNIPARC:UPI0000088A60
R;Joly, E.; Deverson, E.V.; Coadwell, W.L.; Gunther, E.; Howard, J.C.; Butcher, G.W.
submitted to the EMBL Data Library, September 1993
A;Description: The distribution of TAP2 Alleles amongst laboratory RAT RT1 Haplotypes.
A;Reference number: S38402
A;Accession: S38402
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-45,'L',47-703 <JOI>
A;Cross-references: UNIPARC:UPI0000088A60; EMBL:X75307; NID:G407480; PIDN:CAAS3055.1; PI
A;Accession: S38401
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-703 <JOJ>
A;Cross-references: UNIPARC:UPI000008890E; EMBL:X75306; NID:G407478; PIDN:CAAS3054.1; PI
C;Keywords: ATP; nucleotide binding; P-loop
F;486-678/Domain: ATP-binding cassette homology <ABC>
F;503-510/Region: nucleotide-binding motif A (P-loop)

Query Match 5.1%; Score 6; DB 2; Length 703;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQR 92
Db 20 LGLLQR 25
|||||

RESULT 905
G72499
hypothetical protein APE1973 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
A;Accession: G72499
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72499
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-706 <KAW>
A;Cross-references: UNIPROT:Q9YAG6; UNIPARC:UPI000005E173; DDBJ:AP000063; NID:G5105654;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1973
C;Superfamily: Aeropyrum pernix hypothetical protein APE1973

Query Match 5.1%; Score 6; DB 2; Length 706;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db 175 LAALAA 180
|||||

RESULT 906
S61717
Probable membrane protein YOL060c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O1216
C;Species: Saccharomyces cerevisiae
C;Date: 27-Apr-1996 #sequence_revision 10-May-1996 #text_change 09-Jul-2004
A;Accession: S61717; S59287; S66752
R;Mannhaupt, G.; Vetter, I.; Schwarzlose, C.; Mitzel, S.; Feldmann, H.
Yeast 12, 67-76, 1996
A;Title: Analysis of a 26 kb region on the left arm of yeast chromosome XV.
A;Reference number: S61715; MUID:96381248; PMID:8789261
A;Accession: S61717
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-706 <MAN>
A;Cross-references: UNIPROT:Q12296; UNIPARC:UPI000006A7C4; EMBL:X91067; NID:G984177; PID
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
R;Mannhaupt, G.; Vetter, I.; Schwarzlose, C.; Mitzel, S.; Feldmann, H.
submitted to the EMBL Data Library, August 1995
A;Description: Analysis of a 26kb region on the left arm of yeast chromosome XV.
A;Reference number: S59285
A;Accession: S59287
A;Molecule type: DNA
A;Residues: 1-706 <FEI>
A;Cross-references: UNIPARC:UPI000006A7C4; EMBL:X91067; NID:G984177; PID:G984180
R;Feldmann, H.; Mannhaupt, G.; Vetter, I.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66743
A;Accession: S66752
A;Molecule type: DNA
A;Residues: 1-706 <FEW>
A;Cross-references: UNIPARC:UPI000006A7C4; EMBL:Z74802; NID:G1419874; PID:e251867; PID:G
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:AMI3
A;Cross-references: SGD:S0005421
A;Map position: 15L
C;Keywords: transmembrane protein
F;66-82/Domain: transmembrane #status predicted <TM1>
F;151-167/Domain: transmembrane #status predicted <TM2>
F;178-194/Domain: transmembrane #status predicted <TM3>

Query Match 5.1%; Score 6; DB 2; Length 706;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20

Db 72 VLLGGV 77
|||||

RESULT 907

B72619
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 5 APE1411 [similarity] - Aeropyrum F
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B72619
R:Kawarayayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-708 <KAW>
A:Cross-references: UNIPROT:Q9YC40; UNIPARC:UPI000005DF35; DDBJ:AP000061; NID:G5104821;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1411
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5
C:Keywords: membrane-associated complex; NAD; oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 708;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAY 27
|||||

Db 687 AALAY 692
|||||

RESULT 908

E82298
polyribonucleotide nucleotidyltransferase VC0647 [imported] - Vibrio cholerae (strain N1
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82298
R:Heidelber, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Emdalaeva, M.D.; Vamathevan, J.; Bais, S.; Qin, H.; Dragoi, I.; Sellers, E
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82298
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-709 <HEI>
A:Cross-references: UNIPROT:Q9KU76; UNIPARC:UPI00000C2D5C; GB:AE004150; GB:AE003852; NID
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0647
A:Map position: 1
C:Superfamily: polyribonucleotide nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 709;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 QQQAVI 100
|||||

Db 212 QQQAVI 217
|||||

RESULT 909

S23650
retrovirus-related hypothetical protein II - human retrotransposon LINE-1
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 31-Dec-2004
C:Accession: S23650

R:Hohjoh, H.; Minakami, R.; Sakaki, Y.
Nucleic Acids Res. 18, 4099-4104, 1990
A:Title: Selective cloning and sequence analysis of the human L1 (LINE-1) sequences whi
A:Reference number: S23649; MUID:90332398; PMID:2165587
A:Accession: S23650
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-712 <HOH>
A:Cross-references: UNIPROT:Q15604; UNIPROT:Q9UN80; UNIPROT:Q12881; UNIPROT:Q00363; UNI
PROT:Q14754; UNIPROT:Q9Y5K0; UNIPROT:Q00366; UNIPROT:Q8TE30; UNIPROT:Q00375; UNIPARC:UP
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1990
C:Genetics:
A:Mobile element: LINE-1
A:Start codon: GTG

Query Match 5.1%; Score 6; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 9.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 QKLEAF 114
|||||

Db 83 QKLEAF 88
|||||

RESULT 910

T22454
hypothetical protein F49E2.5c - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C:Accession: T22454
R:Sulston, J.
submitted to the EMBL Data Library, October 1994
A:Reference number: Z19566
A:Accession: T22454
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-714 <WIL>
A:Cross-references: UNIPROT:O45541; UNIPARC:UPI0000177522; EMBL:Z46267; PIDN:CAA86424.1;
A:Experimental source: clone F49E2
C:Genetics:
A:Gene: CESP:F49E2.5c
A:Map position: X
A:Introns: 37/2; 66/3; 210/3; 284/3; 314/3; 337/3; 389/3; 531/1; 577/3; 618/3; 655/2; 6

Query Match 5.1%; Score 6; DB 2; Length 714;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 AQVIAH 80
|||||

Db 483 AQVIAH 488
|||||

RESULT 911

A82617
glycyl-tRNA synthetase beta chain XF1959 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: A82617
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number AS9328 below
A:Accession: A82617
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-722 <SIM>
A:Cross-references: UNIPROT:Q9PC26; UNIPARC:UPI00001364AF; GB:AE004015; GB:AE003849; NID
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, P.; Acencio, M.; Alvarenga, R.; P
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, I
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E A;Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A;Gene: XP1959
C;Superfamily: glycine-tRNA ligase beta chain

Query Match 5.1%; Score 6; DB 2; Length 722;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
|||||
Db 451 GKVLGL 456

RESULT 912
S70098
probable membrane protein YDR333c - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C;Accession: S70098
R;Du, Z.
submitted to the EMBL Data Library, March 1996
A;Description: The sequence of S. cerevisiae cosmid 9651.
A;Reference number: S70098
A;Accession: S70098
A;Molecule type: DNA
A;Residues: 1-723 <DUZ>
A;Cross-references: UNIPROT:Q05468; UNIPARC:UPI000006A5CD; EMBL:U51032; NID:gl2330659; PI
C;Genetics:
A;Gene: MIP8;YDR333c
A;Cross-references: SGD:S0002741
A;Map position: 4R
C;Keywords: transmembrane protein
F1620-636/Domain: transmembrane #status predicted <TM>

Query Match 5.1%; Score 6; DB 2; Length 723;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 GLLQRA 93
|||||
Db 390 GLLQRA 395

RESULT 913
WHBEP
infected cell protein ICP18.5 - suid herpesvirus 1 (strain Becker)
C;Species: suid herpesvirus 1
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: A33779
R;Pederson, N.E.; Enquist, L.W.
Nucleic Acids Res. 17, 3597, 1989
A;Title: The nucleotide sequence of a pseudorabies virus gene similar to ICP18.5 of herp
A;Reference number: S04145; MUID:89263808; PMID:2542904
A;Accession: A33779
A;Molecule type: DNA
A;Residues: 1-724 <PED>
A;Cross-references: UNIPROT:P11871; UNIPARC:UPI0000132383; GB:X14573; NID:961407; PIDN:C
C;Superfamily: herpesvirus infected cell protein ICP18.5
C;Keywords: capsid assembly; late protein

Query Match 5.1%; Score 6; DB 1; Length 724;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 236 LAALAA 241

RESULT 914
DB4377
protein export [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: DB4377
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: DB4377
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-724 <STO>
A;Cross-references: UNIPROT:Q9HN37; UNIPARC:UPI0000063ABP; GB:AE004437; NID:gl0581679; P
C;Genetics:
A;Gene: tatC2

Query Match 5.1%; Score 6; DB 2; Length 724;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
|||||
Db 92 GVLAAAL 97

RESULT 915
SS0833
holocarbonylase synthetase - human
C;Species: Homo sapiens (man)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
C;Accession: S50833
R;Suzuki, Y.; Aoki, Y.; Ishida, Y.; Chiba, Y.; Iwamatsu, A.; Kishino, T.; Niikawa, N.; M
Nature Genet. 8, 122-128, 1994
A;Title: Isolation and characterization of mutations in the human holocarbonylase synthe
A;Reference number: S50833; MUID:95144167; PMID:7842009
A;Accession: S50833
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-726 <SUZ>
A;Cross-references: UNIPROT:P50747; UNIPARC:UPI0000126A8C; EMBL:D23672; NID:gs77624; PID
C;Genetics:
A;Gene: GDB:HLCS; HCS
A;Cross-references: GDB:392648; OMIM:253270
A;Map position: 21q22.1-21q22.1
C;Superfamily: Biotin--protein ligase

Query Match 5.1%; Score 6; DB 2; Length 726;
Best Local Similarity 100.0%; Pred. No. 9.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
|||||
Db 241 GKVLGL 246

RESULT 916
DB4754
probable oxidoreductase [EC 1.-.-.-] yagR [similarity] - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: D64754

R; Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A; Title: The complete genome sequence of *Escherichia coli* K-12.
A; Reference number: A64720; MUID: 97426617; PMID: 9278503
A; Accession: D64754
A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA
A; Residues: 1-732 <BLAT>
A; Cross-references: UNIPROT: P77489; UNIPARC: UPI000013A0B1; GB: U00096; NID: 9
A; Experimental source: strain K-12, substrain MG1655
C; Genetics:
A; Gene: *yagR*
C; Superfamily: carbon monoxide dehydrogenase molybdoprotein
C; Keywords: oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 732;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 254 LAALAA 259

RESULT 917
B85518
hypothetical protein *yagR* [imported] - *Escherichia coli* (strain O157:H7, substrain EDL933)
C; Species: *Escherichia coli*
C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C; Accession: B85518
R; Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A; Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A; Reference number: A85480; MUID: 21074935; PMID: 11206551
A; Accession: B85518
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-732 <STO>
A; Cross-references: UNIPROT: Q8X6J4; UNIPARC: UPI000013A0B0; GB: AE005174; NID: g12513067; F
A; Experimental source: strain O157:H7, substrain EDL933
C; Genetics:
A; Gene: *yagR*
C; Superfamily: carbon monoxide dehydrogenase molybdoprotein

Query Match 5.1%; Score 6; DB 2; Length 732;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 254 LAALAA 259

RESULT 918
B90668
hypothetical protein EC0314 [imported] - *Escherichia coli* (strain O157:H7, substrain R
C; Species: *Escherichia coli*
C; Date: 19-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C; Accession: B90668
R; Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A; Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
A; Reference number: A99629; MUID: 21156231; PMID: 11258796
A; Accession: B90668
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-732 <HAY>
A; Cross-references: UNIPROT: Q8X6J4; UNIPARC: UPI000013A0B0; GB: BA000007; PIDN: BAB33737.1;
A; Experimental source: strain O157:H7, substrain RMD 050952
C; Genetics:

A; Gene: EC0314

C; Superfamily: carbon monoxide dehydrogenase molybdoprotein

Query Match 5.1%; Score 6; DB 2; Length 732;

Best Local Similarity 100.0%; Pred. No. 9.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26

|||||

Db 254 LAALAA 259

RESULT 919

F82965

hypothetical protein PA5441 [imported] - *Pseudomonas aeruginosa* (strain PAO1)C; Species: *Pseudomonas aeruginosa*

C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C; Accession: F82965

R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Lim

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A; Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic path

A; Reference number: A82950; MUID: 20437337; PMID: 10984043

A; Accession: F82965

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-733 <STO>

A; Cross-references: UNIPROT: Q9HTC5; UNIPARC: UPI00000C600A; GB: AE004957; GB: AE004091; NI

A; Experimental source: strain PAO1

C; Genetics:

A; Gene: PA5441

Query Match

Best Local Similarity 100.0%; Pred. No. 9.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

|||||

Db 197 LGGVLA 202

RESULT 920

AD2444

hypothetical protein all5108 [imported] - *Nostoc* sp. (strain PCC 7120)C; Species: *Nostoc* sp. PCC 7120A; Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C; Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C; Accession: AD2444

R; Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,

DNA Res. 8, 205-213, 2001

A; Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A; Reference number: AB1807; MUID: 21595285; PMID: 11759840

A; Accession: AD2444

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-733 <KUR>

A; Cross-references: UNIPROT: Q8YM33; UNIPARC: UPI00000CED6F; GB: BA000019; PIDN: BAB76807.1;

A; Experimental source: strain PCC 7120

C; Genetics:

A; Gene: all5108

Query Match

Best Local Similarity 100.0%; Pred. No. 9.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 GHIELG 43

|||||

Db 142 GHIELG 147

RESULT 921

A83006
 Hypothetical protein PA5121 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C:Accession: A83006
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, N.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: A83006
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-735 <STO>
 A:Cross-references: UNIPROT:Q9HU63; UNIPARC:UPI00000C5F11; GB:AE004091; NID:212277337
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA5121

Query Match 5.1%; Score 6; DB 2; Length 735;
 Best Local Similarity 100.0%; Pred. No. 9.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 APYIEQ 74
 |||||
 Db 409 APYIEQ 414

RESULT 922
 D95966
 Probable aldehyde or xanthine dehydrogenase protein [imported] - Sinorhizobium meliloti
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C:Accession: D95966
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan, Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A>Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endosymbiont of the legume Sinorhizobium meliloti
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: D95966
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-741 <KUR>
 A:Cross-references: UNIPROT:Q92U01; UNIPARC:UPI00000CB715; GB:AL591985; PIDN:CAC49396.1
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kies, E.; Komp, C.; Lelaure, habault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Smb21556
 A:Genome: plasmid
 C:Superfamily: carbon monoxide dehydrogenase molybdoprotein

Query Match 5.1%; Score 6; DB 2; Length 741;
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 253 LAALAA 258

RESULT 923
 S61247
 DNA helicase/primase complex associated protein - bovine herpesvirus 1
 C:Species: bovine herpesvirus 1
 C>Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004

C:Accession: S61247
 R:Vicek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Sch submitted to the EMBL Data Library, January 1995
 A:Description: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome
 A:Reference number: S61233
 A:Accession: S61247
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-748 <VLC>
 A:Cross-references: UNIPROT:P52374; UNIPARC:UPI000012C60E; EMBL:Z48053; NID:g971311; PID:212277337
 C:Superfamily: varicella-zoster virus gene 52 protein

Query Match 5.1%; Score 6; DB 2; Length 748;
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
 |||||
 Db 719 LAALAA 724

RESULT 924
 S61643
 Probable membrane protein YOR081c - yeast (Saccharomyces cerevisiae)
 A:Alternate names: hypothetical protein Q2964; hypothetical protein YOR2964c
 C:Species: Saccharomyces cerevisiae
 C>Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004
 C:Accession: S61643; S66966; S66964
 R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia submitted to the EMBL Data Library, December 1995
 A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome 1
 A:Reference number: S61643
 A:Accession: S61643
 A:Molecule type: DNA
 A:Residues: 1-749 <BEN>
 A:Cross-references: UNIPROT:Q12043; UNIPARC:UPI000006B08B; EMBL:X94335; NID:g1262139; PI:212277337
 R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansoerge, W. submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66965
 A:Accession: S66966
 A:Molecule type: DNA
 A:Residues: 1-749 <VOS>
 A:Cross-references: UNIPARC:UPI000006B08B; EMBL:Z74989; NID:g1420242; PID:e2513992; PID:212277337
 A:Experimental source: strain S288C
 R:Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M. submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66929
 A:Accession: S66964
 A:Molecule type: DNA
 A:Residues: 435-749 <BOH>
 A:Cross-references: UNIPARC:UPI0000168EF3; EMBL:Z74989; MIPS:YOR081c
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0005607
 A:Map position: 15R
 C:Keywords: transmembrane protein
 F:211-227/Domain: transmembrane #status predicted <TM1>
 F:327-343/Domain: transmembrane #status predicted <TM2>

Query Match 5.1%; Score 6; DB 2; Length 749;
 Best Local Similarity 100.0%; Pred. No. 9.6e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
 |||||
 Db 197 GVLAAL 202

RESULT 925
 F83003
 Hypothetical protein PA5146 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C;Accession: F83003
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: F83003
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-750 <STO>
 A;Cross-references: UNIPROT:Q9HU38; UNIPARC:UPI000000C5F24; GB:AS004927; GB:AS004091; NID
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA5146

```

Query Match          5.1%; Score 6; DB 2; Length 750;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      84  GKVLGL 89
          | | | | |
Db       5  GKVLGL 10

```

RESULT 926
D72660
probable aldehyde oxidoreductase APE0708 - Aeropyrum pernix (strain K1)
C/Species: Aeropyrum pernix
C/Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C/Accession: D72660
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawarabayashi, Y. 1999
DNA Res. 6, 83-101, 1999
A/Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A/Reference number: A72450; MUID:99310339; PMID:10382966
A/Accession: D72660
A/Molecule type: DNA
A/Residues: 1-753 <KAW>
A/Cross-references: UNIPROT:Q9YE62; UNIPARC:UPI0000005DC61; DDBJ:AP0000060; NID:G5104188;
A/Experimental source: strain K1
C/Genetics:
A/Gene: APE0708
C/Superfamily: carbon monoxide dehydrogenase molybdo-protein

```

Query Match          5.1%; Score 6; DB 1; Length 753;
Best Local Similarity 100.0%; Pred. NO. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
      |||||
Db 746 LAALAA 751

```

RESULT 927
 G87178
 probable DNA-binding protein [imported] - Mycobacterium leprae
 C;Species: Mycobacterium leprae
 C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
 C;Accession: G87178
 R;Cole,S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 N.M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S
 A;Title: Massive gene decay in the leprosy bacillus.
 A;Reference number: A86909; PMID:21128732; PMID:11234002
 A;Accession: G87178
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-753 <STO>
 A;Cross-references: UNIPROT:Q9CBE1; UNIPARC:UPI000000C6E38; GB:AL450380; NID:g13093710; E
 C;Genetics:

A:Gene: ML2156
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0862c

Query Match 5.1%; Score 6; DB 2; Length 753;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRAE 94
|||||
Db 689 LLQRAE 694

RESULT 928
B70838
hypothetical protein Rv0197 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Oct-2004
C:Accession: B70838
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Ho
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.;
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complet
A:Reference number: A70500; MUID:38295987; PMID:9634230
A:Accession: B70838
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-762 <COL>
A:Cross-references: UNIPROT:O53648; UNIPARC:UPI00000C1547; GB:AL021928; GB:AL12
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0197

```

Query Match      5.1%; Score 6; DB 2; Length 762;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 224 LAALAA 229

```

RESULT 929
A47456
down-regulated in adenoma (DRA) - human
C/Species: Homo sapiens (man)
C/Date: 10-May-1996 #sequence_revision 10-May-1996 #text_change 09-Jul-2004
C/Accession: A47456
R/Schweinfest, C.W.; Henderson, K.W.; Suster, S.; Kondoh, N.; Papas, T.S.
Proc. Natl. Acad. Sci. U.S.A. 90, 4166-4170, 1993
A/Title: Identification of a colon mucosa gene that is down-regulated in colon
A/Reference number: A47456; MUID:93248250; PMID:7683425
A/Accession: A47456
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-764 <RES>
A/Cross-references: UNIPROT:P40879; UNIPARC:UPI000000128C; GB:L02785; NID:g2919
C/Genetics:

```

Query Match      5.1%; Score 6; DB 2; Length 764;
Best Local Similarity 100.0%; Pred. No. 9.8e-02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20  VLAALA 25
          |||||
Db      439  VLAALA 444

```

RESULT 930

G96661
Hypothetical protein F24D7.13 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: G96661
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Mafti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, G.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96661
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-767 <STO>
A:Cross-references: UNIPROT:Q9CAD3; UNIPARC:UPI000004838F; GB:AE005173; NID:G6456164; P
C:Genetics:
A:Gene: F24D7.13
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 767;
Best Local Similarity 100.0%; Pred. No. 9.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 20 VLAALA 25
|||||
DB 328 VLAALA 333

RESULT 931

T50299
Hypothetical serine-rich protein [imported] - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jun-2000
C:Accession: T50299
R:Saunders, D.; Harris, D.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 2000
A:Reference number: Z25058
A:Accession: T50299
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: DNA
A:Residues: 1-771 <SAU>
A:Cross-references: UNIPARC:UPI000016905B; EMBL:AL157872; PIDN:CAB75989.1; GSPDB:GN00066
A:Experimental source: strain 972h(-)
C:Genetics:
A:Gene: SPAC6F6.17; SPDB:SPAR736.01
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 771;
Best Local Similarity 100.0%; Pred. No. 9.9e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 47 AIVPDK 52
|||||
DB 537 AIVPDK 542

RESULT 932

T00366
Hypothetical protein KIAA0669 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00366
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete

A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00366
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: mRNA
A:Residues: 1-780 <ISH>
A:Cross-references: UNIPROT:O75157; UNIPARC:UPI00000722E0; EMBL:AB014569; NID:g3327151;
A:Experimental source: brain; clone HK02346
C:Genetics:
A:Note: KIAA0669

Query Match 5.1%; Score 6; DB 2; Length 780;
Best Local Similarity 100.0%; Pred. No. 1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 95 QQQAVI 100
|||||
DB 757 QQQAVI 762

RESULT 933

I47038
vasopressin-activated calcium-mobilizing protein VACM-1 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 09-Jul-2004
C:Accession: I47038; I46537
R:Burnatowska-Hledin, M.A.; Spielman, W.S.; Smith, W.L.; Shi, P.; Meyer, J.M.; Dewitt, D.
Am. J. Physiol. 268, F1198-F1210, 1995
A:Title: Expression cloning of an AVP-activated, calcium-mobilizing receptor from rabbit
A:Reference number: I46537; MUID:95335865; PMID:7611460
A:Accession: I47038
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: mRNA
A:Residues: 1-780 <BUR>
A:Cross-references: UNIPROT:Q29425; UNIPARC:UPI00001380B1; GB:S78157; NID:g1049073; PIDN
A:Accession: I46537
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: mRNA
A:Residues: 1-780 <BU2>
A:Cross-references: UNIPARC:UPI00001380B1; EMBL:U30380; NID:g1049073; PIDN:AAB63562.1; P

Query Match 5.1%; Score 6; DB 2; Length 780;
Best Local Similarity 100.0%; Pred. No. 1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 6 DLEVT 11
|||||
DB 592 DLEVT 597

RESULT 934

T39696
methionyl tRNA synthetase - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39696
R:Wood, V.; Skelton, J.; Churcher, C.M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21870
A:Accession: T39696
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: DNA
A:Residues: 1-782 <WOO>
A:Cross-references: UNIPROT:Q9UUF2; UNIPARC:UPI000006B050; EMBL:AL109652; PIDN:CAB51763
A:Experimental source: strain 972h-, cosmid c17A3
C:Genetics:
A:Gene: pi043
A:Map position: 2
A:Introns: 40/1; 66/2
C:Superfamily: methionyl-tRNA synthetase

Query Match 5.1%; Score 6; DB 2; Length 782;
Best Local Similarity 100.0%; Pred. No. 1e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0

QY 82 FKGVKL 87
|||||
Db 468 FKGVKL 473

RESULT 935

AE2688
Na+/H+ antiporter mnhA [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
A:Accession: AE2688
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.R.; Chen, Y.; Woo, I.
erge, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AE2688
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-789 <KUR>
A:Cross-references: UNIPROT:Q8UGX6; UNIPARC:UPI00000D198D; GB:AE008688; PIDN:AAL41923.1
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: mnhA
A:Map position: circular chromosome

Query Match 5.1%; Score 6; DB 2; Length 789;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 18 LAALAA 23

RESULT 936

H97469
probable NADH dehydrogenase (VCA0157) [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
A:Accession: H97469
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: H97469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-789 <KUR>
A:Cross-references: UNIPROT:Q8UGX6; UNIPARC:UPI00000D198D; GB:AE007869; PIDN:AAK86713.1
C:Genetics:
A:Gene: AGR_C 1658
A:Map position: circular chromosome

Query Match 5.1%; Score 6; DB 2; Length 789;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 18 LAALAA 23

RESULT 937

AB0294
pyruvate, water dikinase (EC 2.7.9.2) [imported] - Yersinia pestis (strain C092)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AB0294
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0294
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-794 <KUR>
A:Cross-references: UNIPROT:Q8ZDY5; UNIPARC:UPI00000DC742; GB:AL590842; PIDN:CAC91214.1
C:Genetics:
A:Gene: ppsA
C:Superfamily: Escherichia coli pyruvate, water dikinase; phosphotransferase system enzy
C:Keywords: transferase

Query Match 5.1%; Score 6; DB 2; Length 794;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GKPAIV 49
|||||
Db 252 GKPAIV 257

RESULT 938

F59433
RhoGAP protein [imported] - human
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 09-Jul-2004
C:Accession: F59433; G59433
R:Richnau, N.; Aspenstrom, P.
J. Biol. Chem. 276, 35060-35070, 2001
A:Title: Rich, a rho gtpase-activating protein domain-containing protein involved in si
A:Reference number: F59433
A:Accession: F59433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <RIC>
A:Cross-references: UNIPROT:Q96KS3; UNIPARC:UPI000004CODE; GB:CAC37948; PID:gl3940243;
R:Aspenstrom, P.
submitted to Genbank, April 2001
A:Reference number: G59433
A:Accession: G59433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-803 <ASP>
A:Cross-references: UNIPARC:UPI000004CODE; GB:CAC37948; PID:gl3940243; PIDN:CAC37948.1

Query Match 5.1%; Score 6; DB 2; Length 803;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 AVIEPI 103
|||||
Db 429 AVIEPI 434

RESULT 939

T24685
hypothetical protein T08D10.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Dec-2004
C:Accession: T24685
R:Lloyd, C.
submitted to the EMBL Data Library, August 1995
A:Reference number: Z19923
A:Accession: T24685
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-803 <WIL>
A:Cross-references: UNIPROT:Q22343; UNIPARC:UPI0000077983; EMBL:Z50756; PIDN:CRAA90637.1

A;Experimental source: clone T08D10

C;Genetics:

A;Gene: CESP:T08D10.2

A;Map position: X

A;Introns: 43/2; 64/1; 135/1; 272/3; 301/1; 345/2; 379/2; 520/1; 586/2; 643/2

Query Match 5.1%; Score 6; DB 2; Length 803;

Best Local Similarity 100.0%; Pred. No. 1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 47 AIVPDK 52

|||||

Db 637 AIVPDK 642

RESULT 940

A12157

hypotheical protein alr2816 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C;Accession: A12157

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: A12157

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-804 <KUR>

A;Cross-references: UNIPROT:O8YTA7; UNIPARC:UPI00000CE585; GB:BA0000019; PIDN:BA074515.1;

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr2816

C;Superfamily: beta-glucosidase, GBA2 type

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 804;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

|||||

Db 601 LAALAA 606

RESULT 941

G87662

peptidase, M20/M25/M40 family [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: G87662

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: AB87249; MUID:21173698; PMID:11259647

A;Accession: G87662

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-805 <STO>

A;Cross-references: UNIPROT:Q9A369; UNIPARC:UPI00000C7A38; GB:AE005673; NID:gl3425037; F

C;Genetics:

A;Gene: CC3337

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 805;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25

|||||

Db 460 VLAALA 465

RESULT 942

B49938

hupU protein - Rhodobacter capsulatus

C;Species: Rhodobacter capsulatus

C;Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 09-Jul-2004

C;Accession: B49938; S24787

R;Elsen, S.; Richaud, P.; Colbeau, A.; Vignais, P.M.

J. Bacteriol. 175, 7404-7412, 1993

A;Title: Sequence analysis and interposon mutagenesis of the hupT gene, which encodes a

A;Reference number: A49938; MUID:94042916; PMID:8226687

A;Accession: B49938

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-806 <ELS>

A;Cross-references: UNIPROT:O86457; UNIPARC:UPI000017AB82; GB:L02348

A;Note: authors translated the codon CTG for residue 30 as Met, TGC for residue 261 as T

R;Richaud, P.

submitted to the EMBL Data Library, January 1991

A;Reference number: S24786

A;Accession: S24787

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-806 <RIC>

A;Cross-references: UNIPARC:UPI000017AB82; EMBL:X57380

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 806;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23

|||||

Db 141 GGVLA 146

RESULT 943

C72858

AcOrf-66 protein - Autographa californica nuclear polyhedrosis virus

C;Species: Autographa californica nuclear polyhedrosis virus, ACMNPV

A;Note: dsDNA virus

C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004

C;Accession: C72858

R;Ayres, M.D.; Howard, S.C.; Kuzio, J.; Lopez-Ferber, M.; Possee, R.D.

Virology 202, 586-605, 1994

A;Title: The complete DNA sequence of Autographa californica nuclear polyhedrosis virus.

A;Reference number: A72850; MUID:94303173; PMID:8030224

A;Accession: C72858

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-808 <AYR>

A;Cross-references: UNIPROT:P41467; UNIPARC:UPI00001391DD; GB:L22858; NID:gs10708; PIDN:

C;Genetics:

A;Gene: AcOrf-66

Query Match

Best Local Similarity 5.1%; Score 6; DB 2; Length 808;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 QOYDEM 62

|||||

Db 725 QOYDEM 730

RESULT 944

T48307

hypotheical protein F9G14.170 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C;Accession: T48307

R;Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De

eves, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, April 2000

A;Reference number: Z24491
A;Accession: T48307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-819 <BEV>
A;Cross-references: UNIPROT:Q9LYZ9; UNIPARC:UPI000000AB55F; EMBL:AL162973
A;Experimental source: cultivar Columbia; BAC clone F9G14
C;Genetics:
A;Map position: 5
A;Note: F9G14..170

Query Match 5.1%; Score 6; DB 2; Length 819;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
DB 530 VLAALA 535

RESULT 945

AE2404
DNA helicase [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AE2404
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata, S.
DNA Rep. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE2404
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-822 <EUR>
A;Cross-references: UNIPROT:Q9MYV5; UNIPARC:UPI000000CEC58; GB:BA0000019; PIDN:BAE76488.1;
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all4789
C;Superfamily: DNA helicase recG

Query Match 5.1%; Score 6; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQR 92
|||||
DB 361 LGLLQR 366

RESULT 946

T29644
hypothetical protein W01A11.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T29644
R;Blanchard, M.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A;Description: the sequence of C. elegans cosmid W01A11.
A;Reference number: Z20658
A;Accession: T29644
A;Status: preliminary;
A;Molecule type: DNA
A;Residues: 1-823 <BLA>
A;Cross-references: UNIPARC:UPI000017BC07; EMBL:U64852; PIDN:AA04966.1; GSPDB:GN00023;
A;Experimental source: strain Bristol N2; clone W01A11
C;Genetics:
A;Gene: CBSP:W01A11.3
A;Map position: 5
A;Introns: 15/3; 87/2; 137/2; 207/1; 258/3; 286/2; 419/3; 594/2; 729/2

Query Match 5.1%; Score 6; DB 2; Length 823;

Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 LAALAA 26
|||||
DB 794 LAALAA 799

RESULT 947

A40894
RNA-directed RNA polymerase (EC 2.7.7.48) - yeast (Saccharomyces cerevisiae) RNA replicase
C;Species: Saccharomyces cerevisiae
C;Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 05-Oct-2004
C;Accession: A40894; A40895
R;Rodriguez-Cousino, N.; Esteban, L.M.; Esteban, R.
J. Biol. Chem. 266, 12772-12778, 1991
A;Title: Molecular cloning and characterization of W double-stranded RNA, a linear molecule
A;Reference number: A40894; MUID:91286317; PMID:20611340
A;Accession: A40894
A;Molecule type: genomic RNA
A;Residues: 1-829 <ROD>
A;Cross-references: UNIPROT:P25328; UNIPARC:UPI0000134B87; GB:M63893; NID:g4090991; PIDN:WATsumoto, Y.; Wickner, R.B.
J. Biol. Chem. 266, 12779-12783, 1991
A;Title: Yeast 20 S RNA replicon. Replication intermediates and encoded putative RNA polymerase
A;Reference number: A40895; MUID:91286318; PMID:1648104
A;Accession: A40895
A;Molecule type: genomic RNA
A;Residues: 1-825, <WAT>
A;Cross-references: UNIPARC:UPI000006BD1D; GB:M64034; NID:g335015; PID:g555371
C;Superfamily: RNA-directed RNA polymerase in W dsRNA
C;Keywords: nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 829;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAA 27
|||||
DB 69 AALAA 74

RESULT 948

S28911
gene DN10 protein - human
C;Species: Homo sapiens (man)
C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C;Accession: S28911; I59566
R;Rinchik, E.M.; Bultman, S.J.; Horsthemke, B.; Lee, S.T.; Strunk, K.M.; Spritz, R.A.; Nature 361, 72-76, 1993
A;Title: A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism
A;Reference number: S28911; MUID:93133287; PMID:8421497
A;Accession: S28911
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-838 <RIN>
A;Cross-references: UNIPROT:Q04671; UNIPARC:UPI0000132E40; GB:M99564; NID:g190284; PIDN:Brilliant, M.H.; King, R.A.; Lyon, M.F.; Lee, S.; Gondo, Y.; Nakatsu, Y.; Gardner, J.J. Science 257, 1121-1124, 1992
A;Title: The Mouse Pink-Eyed Dilution Gene: Association with Human Prader-Willi and Angelman syndromes
A;Reference number: I59566; MUID:92376529; PMID:1509264
A;Accession: I59566
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 288-419 <RES>
A;Cross-references: UNIPARC:UPI0000071770; GB:M97901; NID:g189779; PIDN:AAA36430.1; PIDN:Keywords: transmembrane protein

Query Match 5.1%; Score 6; DB 2; Length 838;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26

```

Db      361 LAALAA 366
|||||
conserved hypothetical protein PA1511 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: C83458
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brady, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: AG2950; MUID:20437337; PMID:10984043
A;Accession: C83458
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-842 <STO>
A;Cross-references: UNIPROT:Q913K1; UNIPARC:UPI00000C53B8; GB:AE004579; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1511

Query Match      5.1%; Score 6; DB 2; Length 842;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      43 GQKPAI 48
|||||
Db      617 GQKPAI 622

RESULT 950
AF3484
hypothetical transmembrane oxidoreductase (EC 1.1.1.-) [imported] - Brucella melitensis
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AF3484
R;DelVecchio, V.G.; Kaputral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: ABJ252; PMID:11756688
A;Accession: AF3484
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-850 <KUR>
A;Cross-references: UNIPROT:Q8YEL8; UNIPARC:UPI0000058227; GB:AE008917; PIDN:AAL53041.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1860
A;Map position: 1
C;Keywords: oxidoreductase

Query Match      5.1%; Score 6; DB 2; Length 850;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      22 AALAA 27
|||||
Db      344 AALAA 349

RESULT 951
HF0939
probable nirX protein - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C;Accession: H70939
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

```

```

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: H70939
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-853 <COL>
A;Cross-references: UNIPROT:O53674; UNIPARC:UPI00000D3AF1; GB:AL021929; GB:AL123456; NID
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: nirX

Query Match      5.1%; Score 6; DB 2; Length 853;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 LLGGVL 21
|||||
Db      387 LLGGVL 392

RESULT 952
A97681
hypothetical protein AGR_C_4860 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: A97681
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: A97681
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-853 <KUR>
A;Cross-references: UNIPROT:Q8UC17; UNIPARC:UPI00000D1FB1; GB:AE007869; PIDN:AAK88402.1;
C;Genetics:
A;Map position: circular chromosome

Query Match      5.1%; Score 6; DB 2; Length 853;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 LAALAA 26
|||||
Db      418 LAALAA 423

RESULT 953
AH2905
conserved hypothetical protein Atu2681 [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AH2905
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
.; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH2905
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-853 <KUR>
A;Cross-references: UNIPROT:Q8UC17; UNIPARC:UPI00000D1FB1; GB:AE008688; PIDN:AAL43662.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:

```

A:Gene: Atu2681

A:Map position: circular chromosome

Query Match 5.1%; Score 6; DB 2; Length 853;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 LAALAA 26
|||||
DB 418 LAALAA 423

RESULT 954

C70740
hypothetical protein Rv1348 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: C70740
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:9825987; PMID:9634230
A:Accession: C70740
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-859 <COL>
A:Cross-references: UNIPROT:Q11018; UNIPARC:UPI000013A849; GB:Z75555; GB:AL123456; NID:9
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1348
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1348; ATP-binding cassette
C:Keywords: ATP
F:626-819/Domain: ATP-binding cassette homology <ABC>

Query Match 5.1%; Score 6; DB 1; Length 859;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 19 GVLAAL 24
|||||
DB 297 GVLAAL 302

RESULT 955

T05470
hypothetical protein T805.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05470
R:Bevan, M.; Medler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15417
A:Accession: T05470
A:Molecule type: DNA
A:Residues: 1-859 <BEV>
A:Cross-references: UNIPROT:Q49711; UNIPARC:UPI00000A87CA; EMBL:AL021890
A:Experimental source: cultivar Columbia; BAC clone T805
C:Genetics:
A:Map position: 4
A:Introns: 43/1; 67/1; 84/2; 304/1; 324/3; 360/3; 398/3; 430/3; 494/3; 534/3; 569/3; 624
A:Note: T805.90

Query Match 5.1%; Score 6; DB 2; Length 859;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 AIVPDK 52
|||||
DB 834 AIVPDK 839

RESULT 956

S24325
glucan 1,4-beta-glucosidase (EC 3.2.1.74) - Pseudomonas fluorescens subsp. cellulosa
N:Alternate names: 1,4-beta-D-glucan glucohydrolase D
C:Species: Pseudomonas fluorescens subsp. cellulosa
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C:Accession: S24325; S35999; S35998
R:Rixon, J.E.; Ferreira, L.M.A.; Durrant, A.J.; Laurie, J.I.; Hazlewood, G.P.; Gilbert, Biochem. J. 285, 947-955, 1992
A:Title: Characterization of the gene celd and its encoded product 1,4-beta-D-glucan glu
A:Reference number: S24325; MUID:92359968; PMID:1497631
A:Accession: S24325
A:Molecule type: DNA
A:Residues: 1-869 <RIX1>
A:Cross-references: UNIPROT:Q47912; UNIPARC:UPI000017A981; EMBL:X65527
A:Accession: S35999
A:Molecule type: Protein
A:Residues: 52-53, 'A', 55-59, 'G', 61-65, 'L', 67 <RIX2>
A:Cross-references: UNIPARC:UPI000017A982
R:Gilbert, H.J.
submitted to the EMBL Data Library, April 1992
A:Reference number: S35998
A:Accession: S35998
A:Molecule type: DNA
A:Residues: 1-26, 'Y', 28-128, 'C', 130-203, 'Y', 205-330, 'LS', 333-374, 'E', 376-473, 'S', 475-80
A:Cross-references: UNIPARC:UPI00000BE452; EMBL:X65527; NID:943408; PIDN:CAA46499.1; PI
C:Genetics:
A:Gene: celd
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 5.1%; Score 6; DB 2; Length 869;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 20 VLAALA 25
|||||
DB 19 VLAALA 24

RESULT 957

AH0172
membrane alanyl aminopeptidase [EC 3.4.11.2] [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AH0172
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AH0172
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-871 <KUR>
A:Cross-references: UNIPROT:Q8ZG92; UNIPARC:UPI00000DCA3D; GB:AL590842; PIDN:CAC90243.1; PI
C:Genetics:
A:Gene: pepN
C:Superfamily: microsomal aminopeptidase
C:Keywords: aminopeptidase

Query Match 5.1%; Score 6; DB 2; Length 871;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 LAALAA 26
|||||
DB 714 LAALAA 719

RESULT 958

AC2417

hypothetical protein alr4891 [imported] - Nostoc sp. (strain PCC 7120)
 C/Species: Nostoc sp. PCC 7120
 A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C/Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C/Accession: AC2417
 R./Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasaomoto, S.; Watanabe, A.; Iriyuchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S. DNA Res. 8, 205-213, 2001
 A/Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena Res. 8, 205-213, 2001
 A/Reference number: AB1807; MUID:21595285; PMID:11759840
 A/Accession: AC2417
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-871 <KUR>
 A/Cross-references: UNIPROT:Q8YMP2; UNIPARC:UPI00000CECAC; GB:BA000019; PIDN:BA076590.1
 A/Experimental source: strain PCC 7120
 C/Genetics:
 A/Gene: alr4891

Query Match 5.1%; Score 6; DB 2; Length 871;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
 |||||
 Db 22 GVLAAL 27

RESULT 959
 AH1058
 Mg(2+) transport ATPase, P-type (EC 3.6.1.-) [imported] - Salmonella enterica subsp. ent
 C/Species: Salmonella enterica subsp. enterica serovar Typhi
 A/Note: this species has also been called Salmonella typhi
 C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 31-Dec-2004
 C/Accession: AH1058
 R./Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P. Nature 413, 848-852, 2001
 A/Authors: Park, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K. A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov A/Reference number: AB0502; MUID:21534947; PMID:11677608
 A/Accession: AH1058
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-902 <PAR>
 A/Cross-references: UNIPARC:UPI000005A978; GB:AL513382; PIDN:CAD06917.1; PID:g16505565;
 C/Genetics:
 A/Gene: mgta
 C/Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom C/Keywords: hydrolase

Query Match 5.1%; Score 6; DB 2; Length 902;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 |||||
 Db 608 LAALAA 613

RESULT 960
 B57147
 Mg2+-transporting ATPase (EC 3.6.1.-) mgta, P-type - Salmonella typhimurium
 C/Species: Salmonella typhimurium
 C/Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 31-Dec-2004
 C/Accession: B57147
 R./Rao, T.; Snaveley, M.D.; Farr, S.G.; Maguire, M.B. J. Bacteriol. 177, 2654-2662, 1995
 A/Title: Magnesium transport in Salmonella typhimurium: mgta encodes a P-type ATPase and A/Reference number: A57147; MUID:95270580; PMID:7751273
 A/Accession: B57147
 A/Status: preliminary

A/Molecule type: DNA
 A/Residues: 1-902 <TAO>
 A/Cross-references: UNIPROT:P36640; UNIPARC:UPI000012628B; GB:U07843; NID:9468205; PIDN: C/Superfamily: Na(+)/K(+)-transporting ATPase alpha chain; ATPase nucleotide-binding dom C/Keywords: hydrolase
 R/543-710/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 5.1%; Score 6; DB 2; Length 902;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 |||||
 Db 608 LAALAA 613

RESULT 961
 T00588
 hypothetical protein At2g30100 [imported] - Arabidopsis thaliana
 N/Alternate names: hypothetical protein T27E13.16
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 C/Accession: T00588; E84704
 R./Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul submitted to the EMBL Data Library, May 1998
 A/Description: Arabidopsis thaliana chromosome II BAC T27E13 genomic sequence.
 A/Reference number: Z14178
 A/Accession: T00588
 A/Status: translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-902 <ROU>
 A/Cross-references: UNIPROT:O64736; UNIPARC:UPI00000A7B5C; EMBL:AC004165; NID:g3150396;
 A/Experimental source: cultivar Columbia
 R./Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L. eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
 A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. A/Reference number: A84420; MUID:20083487; PMID:10617197
 A/Accession: E84704
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-902 <STO>
 A/Cross-references: UNIPARC:UPI00000A7B5C; GB:AE002093; NID:g3150410; PIDN:AAC16962.1; C/Genetics:
 A/Gene: At2g30100; T27E13.16
 A/Map position: 2
 A/Introns: 244/3; 482/3; 581/1; 612/3; 665/3; 702/3; 736/2; 782/2; 820/3; 865/3
 C/Superfamily: Arabidopsis thaliana hypothetical protein T27E13.16; ubiquitin homology F/536-610/Domain: ubiquitin homology <UHH>

Query Match 5.1%; Score 6; DB 2; Length 902;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 PKGKVL 87
 |||||
 Db 579 PKGKVL 584

RESULT 962
 A84212
 hypothetical protein Vng0537c [imported] - Halobacterium sp. NRC-1
 C/Species: Halobacterium sp. NRC-1
 C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
 C/Accession: A84212
 R./Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S. i. Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl Jung, K.H.; Alam, M.; Freitas, T. Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A/Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li A/Title: Genome sequence of Halobacterium species NRC-1.
 A/Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: A84212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-904 <STO>
A:Cross-references: UNIPROT:Q9HRU7; UNIPARC:UPI000006366B; GB:AE004437; NID:g10580137; F82958
C:Genetics: VNG0537C

Query Match 5.1%; Score 6; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 513 LLGGVL 518

RESULT 963
T22457
hypothetical protein F49E2.5d - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22457
R:Sulston, J.
submitted to the EMBL Data Library, October 1994
A:Reference number: Z19566
A:Accession: T22457
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-904 <WIL>
A:Cross-references: UNIPROT:Q45544; UNIPARC:UPI000017B9F0; EMBL:Z46267; PIDN:CAA86427.1;
A:Experimental source: clone F49E2
C:Genetics:
A:Gene: CESP:F49E2.5d
A:Map position: X
A:Introns: 37/2; 66/3; 210/3; 284/3; 367/3; 411/3; 441/3; 527/3; 579/3; 721/1; 76

Query Match 5.1%; Score 6; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 AQVIAH 80
DB 673 AQVIAH 678

RESULT 964
B69435
signal-transducing histidine kinase homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69435
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
-; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69435
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-908 <KLE>
A:Cross-references: UNIPROT:Q28789; UNIPARC:UPI0000056CED; GB:AE001000; GB:AE000782; NID

Query Match 5.1%; Score 6; DB 2; Length 908;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GKPAIV 49
DB 508 GKPAIV 513

A:Accession: A84212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-904 <STO>
A:Cross-references: UNIPROT:Q9HRU7; UNIPARC:UPI000006366B; GB:AE004437; NID:g10580137; F82958
C:Genetics: VNG0537C

Query Match 5.1%; Score 6; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
DB 513 LLGGVL 518

RESULT 963
T22457
hypothetical protein F49E2.5d - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22457
R:Sulston, J.
submitted to the EMBL Data Library, October 1994
A:Reference number: Z19566
A:Accession: T22457
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-904 <WIL>
A:Cross-references: UNIPROT:Q45544; UNIPARC:UPI000017B9F0; EMBL:Z46267; PIDN:CAA86427.1;
A:Experimental source: clone F49E2
C:Genetics:
A:Gene: CESP:F49E2.5d
A:Map position: X
A:Introns: 37/2; 66/3; 210/3; 284/3; 367/3; 411/3; 441/3; 527/3; 579/3; 721/1; 76

Query Match 5.1%; Score 6; DB 2; Length 904;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 AQVIAH 80
DB 673 AQVIAH 678

RESULT 964
B69435
signal-transducing histidine kinase homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: B69435
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
-; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69435
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-908 <KLE>
A:Cross-references: UNIPROT:Q28789; UNIPARC:UPI0000056CED; GB:AE001000; GB:AE000782; NID

Query Match 5.1%; Score 6; DB 2; Length 908;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 GKPAIV 49
DB 508 GKPAIV 513

RESULT 965
F82958
DNA polymerase I PA5493 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: F82958
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
-; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: F82958
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-913 <STO>
A:Cross-references: UNIPROT:Q9HT80; UNIPARC:UPI000006035; GB:AE004962; GB:AE004091; NID
A:Experimental source: strain PAO1
C:Genetics:
A:Gene: POLA; PA5493
C:Superfamily: DNA-directed DNA polymerase I

Query Match 5.1%; Score 6; DB 2; Length 913;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 WQKLEA 113
DB 497 WQKLEA 502

RESULT 966
I48921
DNA ligase (ATP) (EC 6.5.1.1) I - mouse
N:Alternate names: DNA joinase; DNA repair enzyme; polydeoxyribonucleotide synthase (ATP
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I48921; B41275
R:Savini, E.; Biamonti, G.; Ciarocchi, G.; Montecucco, A.
Gene 144, 253-257, 1994
A:Title: Cloning and sequence analysis of a cDNA coding for the murine DNA ligase I enzy
A:Reference number: I48921; MUID:94314225; PMID:8039710
A:Accession: I48921
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-916 <RES>
A:Cross-references: UNIPROT:P37913; UNIPARC:UPI00000299F2; EMBL:U04674; NID:g905353; PFI
R:Petrini, J.H.J.; Huwiler, K.G.; Weaver, D.T.
Proc. Natl. Acad. Sci. U.S.A. 88, 7615-7619, 1991
A:Title: A wild-type DNA ligase I gene is expressed in Bloom's syndrome cells.
A:Reference number: A41275; MUID:91352039; PMID:1881902
A:Accession: B41275
A:Molecule type: mRNA
A:Residues: 'NK', 716-746, 'IRT', <PET>
A:Cross-references: UNIPARC:UPI000017629E
C:Superfamily: yeast polydeoxyribonucleotide synthase (ATP)
C:Keywords: DNA repair; ligase; phosphoprotein
F:566/Active site: Lys (covalent AMP-binding) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 916;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
DB 457 VLAALA 462

RESULT 967
T47846
Arm repeat containing protein-like - Arabidopsis thaliana

N:Alternate names: protein T8B10.10
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: T47846
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Mayer, submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24478
A:Accession: T47846
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-928 <RIS>
A:Cross-references: UNIPROT:Q9M224; UNIPARC:UPI00000A6752; EMBL:AL1138646
A:Experimental source: cultivar Columbia; BAC clone T8B10
C:Genetics:
A:Map position: 3
A:Map position: 504/3; 531/3; 575/3; 647/3; 679/2; 700/3; 759/2; 861/3
A:Introns: T8B10.10
A:Note: T8B10.10

Query Match 5.1%; Score 6; DB 2; Length 928;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PKGKVL 87
|||||
Db 302 PKGKVL 307

RESULT 968
T34206
hypothetical protein F10E7.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34206
R:Pauley, A.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F10E7.
A:Reference number: Z21489
A:Accession: T34206
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-929 <PAU>
A:Cross-references: UNIPROT:Q19300; UNIPARC:UPI0000082128; EMBL:U41264; PIDN:AAA82421.1;
C:Genetics:
A:Gene: CESP:F10E7.8
A:Introns: 13/1; 70/3; 120/3; 180/3; 672/3; 701/1; 855/1

Query Match 5.1%; Score 6; DB 2; Length 929;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SADLEV 9
|||||
Db 102 SADLEV 107

RESULT 969
T00403
hypothetical protein At2g44900 [imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T13E15.9
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C:Accession: T00403; B84884
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Maso submitted to the EMBL Data Library, July 1997
A:Description: Arabidopsis thaliana chromosome II BAC T13E15 genomic sequence.
A:Reference number: Z14146
A:Accession: T00403
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-930 <ROU>
A:Cross-references: UNIPROT:Q22161; UNIPARC:UPI00000A20F7; EMBL:AC002388; NID:g3420042;
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L. euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84884
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-930 <STO>
A:Cross-references: UNIPARC:UPI00000A20F7; GB:AE002093; NID:g2344894; PIDN:AAC31834.1;
C:Genetics:
A:Gene: At2g44900; T13E15.9
A:Map position: 2
A:Map position: 460/3; 513/3; 540/3; 584/3; 656/3; 688/2; 709/3; 768/2; 844/2; 872/3
A:Introns: 460/3; 513/3; 540/3; 584/3; 656/3; 688/2; 709/3; 768/2; 844/2; 872/3

Query Match 5.1%; Score 6; DB 2; Length 930;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 82 PKGKVL 87
|||||
Db 310 PKGKVL 315

RESULT 970
T39624
6-phosphofructokinase beta subunit - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39624
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z21843
A:Accession: T39624
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-942 <WOO>
A:Cross-references: UNIPROT:Q42938; UNIPARC:UPI00000697AB; EMBL:AL022104; PIDN:CAA17900.
A:Experimental source: strain 972h-; cosmid c16H5
C:Genetics:
A:Gene: SPDB:SPBC16H5.02
A:Map position: 2
C:Superfamily: ATP-dependent phosphofructokinase, eukaryotic type; 6-phosphofructokinase

Query Match 5.1%; Score 6; DB 2; Length 942;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 396 VLAALA 401

RESULT 971
B90994
probable molybdate metabolism regulator [imported] - Escherichia coli (strain O157:H7, s
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: B90994
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A:Reference number: A39629; MUID:21156231; PMID:11258796
A:Accession: B90994
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-947 <HAY>
A:Cross-references: UNIPROT:Q8X7E5; UNIPARC:UPI00001654A2; GB:BA000007; PIDN:BA836345.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: EC82922

Query Match 5.1%; Score 6; DB 2; Length 947;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 451 VLAALA 456

RESULT 972
B83200
probable phosphotransferase system enzyme I PA3562 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: B83200
R:Stover, C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Boman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83200
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-956 <STO>
A:Cross-references: UNIPROT:Q9HY55; UNIPARC:UPI000000C5A6C; GB:AE004777; GB:AE004091; NID: T22459
C:Genetics:
A:Gene: PA3562
C:Superfamily: fructose phosphotransferase multiphosphoryltransfer protein; phosphotransferase system phosphohistidine-containing protein homology

Query Match 5.1%; Score 6; DB 2; Length 956;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGLLQ 91
|||||
Db 135 VLGLLQ 140

RESULT 973
A49847
nitrite reductase [NAD(P)H] (EC 1.7.1.4) nasB - Klebsiella pneumoniae
N:Alternate names: assimilatory nitrite reductase
C:Species: Klebsiella pneumoniae
C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 03-Jun-2002
C:Accession: A49847
R:Lin, J.T.; Goldman, B.S.; Stewart, V.
J. Bacteriol. 175, 2370-2378, 1993
A:Title: Structures of genes nasA and nasB, encoding assimilatory nitrate and nitrite reductase
A:Reference number: A49847; MUID:93224460; PMID:8468296
A:Accession: A49847
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-957 <LIN>
A:Cross-references: UNIPARC:UPI000017CD9C
A:Experimental source: M5al
A:Note: sequence extracted from NCBI backbone (NCBIP:129204)
C:Keywords: oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 957;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVL 21
|||||
Db 366 LLGGVL 371

RESULT 974
T22459
hypothetical protein F49E2.5b - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22459
R:Sulston, J.
submitted to the EMBL Data Library, October 1994
A:Reference number: Z19566
A:Accession: T22459
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-962 <WIL>
A:Cross-references: UNIPROT:Q20625; UNIPARC:UPI000017B9BF; EMBL:Z46267; PIDN:CAA86429.1;
A:Experimental source: clone F49E2
C:Genetics:
A:Gene: CESP:F49E2.5b
A:Map position: X
A:Introns: 37/2; 66/3; 210/3; 284/3; 469/3; 499/3; 522/3; 585/3; 637/3; 779/1; 825/3; 8;

Query Match 5.1%; Score 6; DB 2; Length 962;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 AQVIAH 80
|||||
Db 731 AQVIAH 736

RESULT 975
B85839
probable regulator molR_B molR_B [imported] - Escherichia coli (strain O157:H7, substra

C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: B85839
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, L.; Fink, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B85839
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-963 <STO>
A:Cross-references: UNIPROT:Q8X7E5; UNIPARC:UPI000000D5E07; GB:AE005174; NID:gl2516332; 8;

Query Match 5.1%; Score 6; DB 2; Length 963;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALA 25
|||||
Db 467 VLAALA 472

RESULT 976
G87389
TonB-dependent receptor [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87389
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolor, n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87389
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-976 <STO>
A:Cross-references: UNIPROT:Q9A963; UNIPARC:UPI000000C72A5; GB:AE005673; NID:gl13422445; F

A:Gene: CC1131

Query Match 5.1%; Score 6; DB 2; Length 976;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
|||||
DB 325 GVLAAL 330

RESULT 977

T34830
polyketide synthase redX - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C:Accession: T34830
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z21558
A:Accession: T34830
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-982 <OLI>
A:Cross-references: UNIPROT:054142; UNIPARC:UPI00000DABP0; EMBL:AL021530; PIDN:CAA16487.
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: redX; SCOEDB:SC2E9.19

Query Match 5.1%; Score 6; DB 2; Length 982;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 663 LAALAA 668

RESULT 978

T39902
translation elongation factor 2 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T39902
R:Llyne, M.; Rajandream, M.A.; Barrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, November 1998
A:Reference number: Z21889
A:Accession: T39902
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-983 <LYN>
A:Cross-references: UNIPROT:094316; UNIPARC:UPI0000069626; EMBL:AL033534; PIDN:CAA22126.
A:Experimental source: strain 972h-; cosmid c215
C:Genetics:
A:Gene: SPDB:SPBC215.12
A:Map position: 2
A:Introns: 8/2; 128/2; 245/3
C:Superfamily: translation elongation factor 2; translation elongation factor Tu homolog

Query Match 5.1%; Score 6; DB 2; Length 983;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
|||||
DB 563 VLLGGV 568

RESULT 979

GB3022
probable two-component sensor PA4982 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C:Accession: GB3022
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: GB3022
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-998 <STO>
A:Cross-references: UNIPROT:Q9HUI3; UNIPARC:UPI00000C5EAB; GB:AE004911; GB:AE004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4982

Query Match 5.1%; Score 6; DB 2; Length 998;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 201 LAALAA 206

RESULT 980

AB3467
sarcosine oxidase (EC 1.5.3.1) [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AB3467
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova, J.; Mazur, M.; Golteaman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesee
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3467
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1000 <KUR>
A:Cross-references: UNIPROT:Q8VP07; UNIPARC:UPI000005819C; GB:AE008917; PIDN:AAL52901.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11720
A:Map position: 1
C:Keywords: oxidoreductase

Query Match 5.1%; Score 6; DB 2; Length 1000;

Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
DB 230 LAALAA 235

RESULT 981

JC5062
phogrin precursor - human
N:Contains: protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type
C:Species: Homo sapiens (man)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C:Accession: JC5062; T46903
R:Kawasaki, E.; Hutton, J.C.; Eisenbarth, G.S.
Biochem. Biophys. Res. Commun. 227, 440-447, 1996
A:Title: Molecular cloning and characterization of the human transmembrane protein tyros
A:Reference number: JC5062; MUID:97032784; PMID:8878534
A:Contents: 1slet
A:Accession: JC5062
A:Molecule type: mRNA
A:Residues: 1-1015 <KAW>
A:Cross-references: UNIPROT:Q92932; UNIPARC:UPI00001329A8; GB:U66702; NID:g1620663; PIDN
R:Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.

RESULT 982

JC5062
phogrin precursor - human
N:Contains: protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type
C:Species: Homo sapiens (man)
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 09-Jul-2004
C:Accession: JC5062; T46903
R:Kawasaki, E.; Hutton, J.C.; Eisenbarth, G.S.
Biochem. Biophys. Res. Commun. 227, 440-447, 1996
A:Title: Molecular cloning and characterization of the human transmembrane protein tyros
A:Reference number: JC5062; MUID:97032784; PMID:8878534
A:Contents: 1slet
A:Accession: JC5062
A:Molecule type: mRNA
A:Residues: 1-1015 <KAW>
A:Cross-references: UNIPROT:Q92932; UNIPARC:UPI00001329A8; GB:U66702; NID:g1620663; PIDN
R:Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.

Biochem. Biophys. Res. Commun. 229, 402-411, 1996
A;Title: ICAAR, a novel member of a new family of transmembrane, tyrosine phosphatase-11
A;Reference number: JCS263; MUID:97127415; PMID:8954911
A;Accession: JCS263
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-207, 'S', 209-246, 'G', 248-322, 'R', 324, 'N', 326-1015 <SMI>
A;Cross-references: UNIPARC:UPI000016AADG; GB:Y08569; NID:91844377; PIDN:CAA69880.1; PID
R;Ansoorge, W.; Winkler, U.; Mewes, H.W.; Well, B.; Wiemann, S.
submitted to the Protein Sequence Database, February 2000
A;Reference number: Z24134
A;Accession: T46903
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: DA', 714, 771-1015 <AAA>
A;Cross-references: UNIPARC:UPI00000725AF; EMBL:AL157451
A;Experimental source: adult amygdala; clone DKFZp761A0712
C;Comment: This protein has an intracellular protein tyrosine phosphatase like protein.
C;Genetics:
A;Note: DKFZp761A0712.1
C;Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphatase
C;Keywords: phosphoprotein; phosphoric monoester hydrolase; transmembrane protein; tyrosine
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-1015/Product: phogrin #status predicted <MAT>
F;615-639/Domain: transmembrane #status predicted <TMM>
F;770-994/Domain: protein-tyrosine-phosphatase homology <PTP2>
F;945/Active site: Cys (phosphotyrosine intermediate) #status predicted
F;951/Binding site: substrate phosphate (Arg) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 1015;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCVVIV 37
|||||
Db 838 GCVVIV 843

RESULT 982
JCS263
transmembrane tyrosine phosphatase-like protein, ICAAR - human
C;Species: Homo sapiens (man)
C;Date: 25-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C;Accession: JCS263
R;Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.
Biochem. Biophys. Res. Commun. 229, 402-411, 1996
A;Title: ICAAR, a novel member of a new family of transmembrane, tyrosine phosphatase-11
A;Reference number: JCS263; MUID:97127415; PMID:8954911
A;Accession: JCS263
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1015 <SMI>
A;Cross-references: UNIPARC:UPI000016AADG; GB:Y08569; NID:91844377; PIDN
C;Comment: This protein has an intracellular protein tyrosine phosphatase like protein.
C;Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphatase
C;Keywords: phosphoprotein
F;770-994/Domain: protein-tyrosine-phosphatase homology <PTP2>
F;945/Active site: Cys (phosphotyrosine intermediate) #status predicted
F;951/Binding site: substrate phosphate (Arg) #status predicted

Query Match 5.1%; Score 6; DB 2; Length 1015;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCVVIV 37
|||||
Db 838 GCVVIV 843

RESULT 983
T30542
major surface glycoprotein - Pneumocystis carinii (fragment)
C;Species: Pneumocystis carinii

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 15-Jun-2001
C;Accession: T30542
R;Mei, Q.; Turner, R.E.; Sorial, V.; Klivington, D.; Angus, C.W.; Kovacs, J.A.
Infect. Immun. 66, 4268-4273, 1998
A;Title: Characterization of major surface glycoprotein genes of human Pneumocystis car
A;Reference number: Z17905; MUID:98380374; PMID:9712777
A;Accession: T30542
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1017 <MEI>
A;Cross-references: UNIPARC:UPI000006A3F8; EMBL:AF033210; NID:g3560516; PID:g3560517; P
C;Genetics:
A;Gene: MSG
C;Superfamily: Pneumocystis carinii major surface glycoprotein MSG100

Query Match 5.1%; Score 6; DB 2; Length 1017;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
|||||
Db 245 IELGGK 250

RESULT 984
A83613
conserved hypothetical protein PA0262 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: A83613
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: A83613
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1019 <STO>
A;Cross-references: UNIPROT:Q916W7; UNIPARC:UPI000000C4FCB; GB:AE004464; GB:AE004091; NI
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA0262

Query Match 5.1%; Score 6; DB 2; Length 1019;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 GKKPAI 48
|||||
Db 617 GKKPAI 622

RESULT 985
H75423
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: H75423
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75423
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1021 <WHI>
A;Cross-references: UNIPROT:Q9RV29; UNIPARC:UPI000000C18C4; GB:AE001969; GB:AE000513; NI
A;Experimental source: strain R1
C;Genetics:

A:Gene: DR1201
A:Map position: 1

Query Match 5.1%; Score 6; DB 2; Length 1021;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
|||||
Db 456 LAALAA 461

RESULT 986

T30543
major surface glycoprotein - Pneumocystis carinii (fragment)
C:Species: Pneumocystis carinii
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 15-Jun-2001
C:Accession: T30543
R:Mei, Q.; Turner, R.E.; Sorial, V.; Klivington, D.; Angus, C.W.; Kovacs, J.A.
Infest. Immun. 66, 4288-4273, 1998
A>Title: Characterization of major surface glycoprotein genes of human Pneumocystis carinii
A:Reference number: 217905; MUID:96380374; PMID:9712777
A:Accession: T30543
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1022 <MEI>
A:Cross-references: UNIPARC:UPI0000069846; EMBL:AF033211; NID:g3560518; PID:g3560519; PI0000069846
C:Genetics:
A:Gene: MSG
C:Superfamily: Pneumocystis carinii major surface glycoprotein MSG100

Query Match 5.1%; Score 6; DB 2; Length 1022;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGK 45
|||||
Db 245 IELGGK 250

RESULT 987

F96602
hypothetical protein T6H22.8.2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2004
C:Accession: F96602
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, T.H.; Dewar, K.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96602
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1029 <STO>
A:Cross-references: UNIPROT:Q9SGT9; UNIPARC:UPI00000A0114; GB:AE005173; NID:g6056374; PI00000A0114
C:Genetics:
A:Gene: T6H22.8.2
A:Map position: 1
C:Superfamily: Receptor-like protein kinase

Query Match 5.1%; Score 6; DB 2; Length 1029;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 10
|||||

Db 211 ADLEVT 216

RESULT 988

C97665
multidrug efflux transporter rmd family protein (AF232237) [imported] - Agrobacterium tumefaciens
C:Species: Agrobacterium tumefaciens
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: C97665
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A>Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: C97665
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1029 <KUR>
A:Cross-references: UNIPROT:Q8UCE5; UNIPARC:UPI00000D1F3D; GB:AE007869; PIDN:AAK88276.1; C:Genetics:
A:Gene: AGR_C_4621
A:Map position: circular chromosome
C:Superfamily: acriflavin resistance protein

Query Match 5.1%; Score 6; DB 2; Length 1029;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 21 LGGVLA 26

RESULT 989

AF2889
acriflavin resistance protein B [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AF2889
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.; erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Ito, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF2889
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1029 <KUR>
A:Cross-references: UNIPROT:Q8UCE5; UNIPARC:UPI00000D1F3D; GB:AE008688; PIDN:AAI43532.1; A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: acrB
A:Map position: circular chromosome
C:Superfamily: acriflavin resistance protein

Query Match 5.1%; Score 6; DB 2; Length 1029;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
|||||
Db 21 LGGVLA 26

RESULT 990

T30331
P-glycoprotein - Trypanosoma cruzi
C:Species: Trypanosoma cruzi
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30331

R.Barreiro, L.; Gamarro, F.; Castany, S.
submitted to the EMBL Data Library, March 1997
A:Description: Tcpgpla, a P-glycoprotein gene of Trypanosoma cruzi.
A:Reference number: Z20824
A:Accession: T30331
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1034 <BAR>
A:Cross-references: UNIPROT:Q9U097; UNIPARC:UPI00000788C3; EMBL:U95956; NID:g3004631; PID:
C:Genetics:
A:Gene: Dgpla

Query Match 5.1%; Score 6; DB 2; Length 1034;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 344 LAALAA 349
|||||
RESULT 991
T31097
chitin synthase (EC 2.4.1.16) CHS1 - fungus (Filobasidium floriforme)
C:Species: Filobasidiella neoformans, Cryptococcus neoformans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
A:Accession: T31097
R:Specht, C.A.
submitted to the EMBL Data Library, August 1997
A:Description: Chs1, a class IV chitin synthase of Cryptococcus neoformans.
A:Reference number: Z20980
A:Accession: T31097
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1041 <SPE>
A:Cross-references: UNIPROT:O13356; UNIPARC:UPI00001278CD; EMBL:AF021318; NID:g2444456;
A:Experimental source: strain H99
C:Genetics:
A:Gene: CHS1

A:Introns: 152/1; 556/1; 634/3; 922/2
C:Function:
A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucose
A:Pathway: chitin biosynthesis
A:Note: class IV chitin synthase
C:Superfamily: chitin synthase chs4
C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 5.1%; Score 6; DB 2; Length 1041;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
DB 116 VLLGGV 121
|||||

RESULT 992
B27672
RNA-directed DNA polymerase (EC 2.7.7.49) homolog (R1) - silkworm
N:Alternate names: reverse transcriptase homolog (R1)
C:Species: Bombyx mori (silkworm)
C:Date: 20-Jun-1989 #sequence_revision 20-Jun-1989 #text_change 09-Jul-2004
A:Accession: B27672
R:Xiong, Y.; Eickbush, T.H.
Mol. Cell. Biol. 8, 114-123, 1988
A:Title: The site-specific ribosomal DNA insertion element R1b belongs to a class of no
A:Reference number: A93098; MUID:88094376; PMID:2447482
A:Accession: B27672
A:Molecule type: DNA
A:Residues: 1-1051 <XIO>
A:Cross-references: UNIPROT:Q7M474; UNIPARC:UPI0000179233
C:Superfamily: silkworm pol protein
C:Keywords: nucleotidyltransferase

Query Match 5.1%; Score 6; DB 2; Length 1034;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
DB 344 LAALAA 349
|||||
RESULT 991
T31097
chitin synthase (EC 2.4.1.16) CHS1 - fungus (Filobasidium floriforme)
C:Species: Filobasidiella neoformans, Cryptococcus neoformans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
A:Accession: T31097
R:Specht, C.A.
submitted to the EMBL Data Library, August 1997
A:Description: Chs1, a class IV chitin synthase of Cryptococcus neoformans.
A:Reference number: Z20980
A:Accession: T31097
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1041 <SPE>
A:Cross-references: UNIPROT:O13356; UNIPARC:UPI00001278CD; EMBL:AF021318; NID:g2444456;
A:Experimental source: strain H99
C:Genetics:
A:Gene: CHS1

Query Match 5.1%; Score 6; DB 2; Length 1051;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 10
DB 862 ADLEVT 867
|||||

RESULT 993
S58883
calcium-channel homolog - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
A:Accession: S58883
R:Schafer, W.R.; Kenyon, C.J.
Nature 375, 73-78, 1995
A:Title: A calcium-channel homolog required for adaptation to dopamine and serotonin
A:Reference number: S58883; MUID:95240736; PMID:7723846
A:Accession: S58883
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1053 <SCH>
A:Cross-references: UNIPROT:Q17354; UNIPARC:UPI0000077480; EMBL:U25119; NID:g841433; PID:
C:Genetics:
A:Introns: 43/1; 185/3; 226/1; 297/1; 392/3; 428/2; 463/3; 724/3; 942/2
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 5.1%; Score 6; DB 2; Length 1053;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 CVVIVG 38
DB 79 CVVIVG 84
|||||

RESULT 994
S09112
hypothetical protein 3 - fruit fly (Drosophila melanogaster) transposable element R2
C:Species: Drosophila melanogaster
C:Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
A:Accession: S09112
R:Jakubczak, J.L.; Xiong, Y.; Eickbush, T.H.
J. Mol. Biol. 212, 37-52, 1990
A:Title: Type I (R1) and type II (R2) ribosomal DNA insertions of Drosophila melanogaster
A:Reference number: S09110; MUID:90204545; PMID:1690812
A:Accession: S09112
A:Molecule type: DNA
A:Residues: 1-1057 <JAK>
A:Cross-references: UNIPROT:P16423; UNIPARC:UPI0000131EAS; EMBL:X51967; NID:g8432; PID:
A:Note: this reading frame extends between two stop codons and does not begin with a start
C:Genetics:
A:Gene: FlyBase:R2-element
A:Cross-references: FlyBase:FBgn0003909
A:Mobile element: transposable element R2

Query Match 5.1%; Score 6; DB 2; Length 1057;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 GCWVIV 37
DB 924 GCWVIV 929
|||||

RESULT 995
H97657
ceoB protein [imported] - Agrobacterium tumefaciens (strain C58, Cerson)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: H97657

A:Reference number: S65154
 A:Accession: S65171
 A:Molecule type: DNA
 A:Residues: 1-1090 <PUR>
 A:Cross-references: UNIPARC:UPI0000052EC7; EMBL:Z73516; NID:gl1370339; PIDN:CAA97865.1; R
 A:Experimental source: strain S288C (AB972)
 R:Purnelle, B.; Comblez, S.; Coster, F.; Naveau, F.; Goffeau, A.
 submitted to the EMBL Data Library, March 1996
 A:Description: The sequence of 55 kb on the left arm of yeast chromosome XVI identifies
 ogue to the human phosphotyrosyl phosphatase activator PTPA and a homologue to the plant
 A:Reference number: S69428
 A:Accession: S69439
 A:Molecule type: DNA
 A:Residues: 1-1090 <PUW>
 A:Cross-references: UNIPARC:UPI0000052EC7; EMBL:X96770; NID:gl1403537; PIDN:CAA65561.1; R
 C:Genetics:
 A:Gene: SGD:CDC60
 A:Cross-references: SGD:S0006081; MIPS:YPL160W
 A:Map position: 16L
 C:Superfamily: Neurospora cytosolic leucine-trRNA ligase
 C:Keywords: aminoacyl-trRNA synthetase; cytosol; ligase; protein biosynthesis

Query Match 5.1%; Score 6; DB 2; Length 1090;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
 |||||
 DB 907 GVLAAAL 912

RESULT 1000
 AF2953
 DNA polymerase III, alpha chain [imported] - Agrobacterium tumefaciens (strain C58, Dupo
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C:Accession: AF2953
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; MUID:21608550; PMID:11743193
 A:Accession: AF2953
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1091 <KUR>
 A:Cross-references: UNIPROT:Q8UAY9; UNIPARC:UPI00000D2117; GB:AE008689; PIDN:AAL44044.1;
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: dnaE
 A:Map position: linear chromosome

Query Match 5.1%; Score 6; DB 2; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
 |||||
 DB 421 GKVLGL 426

Search completed: January 27, 2006, 19:30:01
 Job time : 56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: January 27, 2006, 19:25:16 ; Search time 161 Seconds
(without alignments)
517.095 Million cell updates/sec

Title: US-09-638-693A-36_COPY_16_133
Perfect score: 118
Sequence: 1 ACMSADLEVTSTWLLGV.....VIEPIVTTNQKLEAFWVHKH 118

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 2166443 seqs, 705528306 residues

Word size : 0
Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 2000 summaries

Database : UniProt.05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	83	70.3	209	2 Q81594_9HEPC
2	47	39.8	133	2 Q81595_9HEPC
3	47	39.8	3021	2 Q92933_9HEPC
4	47	39.8	3021	2 Q6870_9HEPC
5	42	35.6	3021	2 Q81258_9HEPC
6	35	29.7	70	2 Q38899_9HEPC
7	35	29.7	88	2 Q38898_9HEPC
8	35	29.7	90	2 Q39904_9HEPC
9	35	29.7	90	2 Q39906_9HEPC
10	35	29.7	90	2 Q39908_9HEPC
11	35	29.7	90	2 Q39910_9HEPC
12	35	29.7	138	2 Q68233_9HEPC
13	32	27.1	83	2 Q39900_9HEPC
14	31	26.3	84	2 Q39896_9HEPC
15	31	26.3	84	2 Q39917_9HEPC
16	30	25.4	3023	2 Q81487_9HEPC
17	29	24.6	89	2 Q39891_9HEPC
18	29	24.6	89	2 Q39894_9HEPC
19	29	24.6	90	2 Q39905_9HEPC
20	28	23.7	84	2 Q39901_9HEPC
21	28	23.7	90	2 Q39902_9HEPC
22	28	23.7	90	2 Q39912_9HEPC
23	28	23.7	90	2 Q39913_9HEPC
24	28	23.7	90	2 Q39916_9HEPC
25	28	23.7	90	2 Q68233_9HEPC
26	23	19.5	138	2 Q68241_9HEPC
27	23	19.5	138	2 Q81495_9HEPC
28	23	19.5	3021	2 Q39909_9HEPC
29	22	18.6	90	2 Q39907_9HEPC
30	21	17.8	87	2 Q56637_9HEPC
31	20	16.9	193	2 Q81594_9HEPC
				Q81595_9HEPC
				Q92933_9HEPC
				Q6870_9HEPC
				Q81258_9HEPC
				Q38899_9HEPC
				Q38898_9HEPC
				Q39904_9HEPC
				Q39906_9HEPC
				Q39908_9HEPC
				Q39910_9HEPC
				Q68233_9HEPC
				Q39900_9HEPC
				Q39896_9HEPC
				Q39917_9HEPC
				Q81487_9HEPC
				Q39891_9HEPC
				Q39894_9HEPC
				Q39905_9HEPC
				Q39901_9HEPC
				Q39902_9HEPC
				Q39912_9HEPC
				Q39913_9HEPC
				Q39916_9HEPC
				Q68233_9HEPC
				Q68241_9HEPC
				Q81495_9HEPC
				Q39909_9HEPC
				Q39907_9HEPC
				Q56637_9HEPC

32	19	16.1	3015	2 Q4QTD9_9HEPC	Q4QTD9_9HEPC
33	19	16.1	3018	2 Q39927_9HEPC	Q39927_9HEPC
34	19	16.1	3019	2 Q512N3_9HEPC	Q512N3_9HEPC
35	18	15.3	3016	2 Q92531_9HEPC	Q92531_9HEPC
36	18	15.3	3019	2 Q68801_9HEPC	Q68801_9HEPC
37	15	12.7	3008	2 Q39929_9HEPC	Q39929_9HEPC
38	15	12.7	3015	2 Q92532_9HEPC	Q92532_9HEPC
39	13	11.0	138	2 Q68222_9HEPC	Q68222_9HEPC
40	13	11.0	138	2 Q68223_9HEPC	Q68223_9HEPC
41	13	11.0	138	2 Q68224_9HEPC	Q68224_9HEPC
42	13	11.0	138	2 Q68225_9HEPC	Q68225_9HEPC
43	13	11.0	138	2 Q68226_9HEPC	Q68226_9HEPC
44	13	11.0	138	2 Q68236_9HEPC	Q68236_9HEPC
45	13	11.0	138	2 Q68243_9HEPC	Q68243_9HEPC
46	13	11.0	313	2 Q02258_9HEPC	Q02258_9HEPC
47	13	11.0	636	2 Q68K64_9HEPC	Q68K64_9HEPC
48	13	11.0	652	2 Q68K51_9HEPC	Q68K51_9HEPC
49	13	11.0	658	2 Q68K66_9HEPC	Q68K66_9HEPC
50	13	11.0	658	2 Q68K65_9HEPC	Q68K65_9HEPC
51	13	11.0	658	2 Q68K61_9HEPC	Q68K61_9HEPC
52	13	11.0	658	2 Q68K34_9HEPC	Q68K34_9HEPC
53	13	11.0	659	2 Q68K67_9HEPC	Q68K67_9HEPC
54	13	11.0	659	2 Q68K62_9HEPC	Q68K62_9HEPC
55	13	11.0	659	2 Q68K60_9HEPC	Q68K60_9HEPC
56	13	11.0	659	2 Q68K59_9HEPC	Q68K59_9HEPC
57	13	11.0	659	2 Q68K56_9HEPC	Q68K56_9HEPC
58	13	11.0	659	2 Q68K55_9HEPC	Q68K55_9HEPC
59	13	11.0	659	2 Q68K54_9HEPC	Q68K54_9HEPC
60	13	11.0	659	2 Q68K49_9HEPC	Q68K49_9HEPC
61	13	11.0	659	2 Q68K47_9HEPC	Q68K47_9HEPC
62	13	11.0	659	2 Q68K45_9HEPC	Q68K45_9HEPC
63	13	11.0	659	2 Q68K43_9HEPC	Q68K43_9HEPC
64	13	11.0	659	2 Q68K42_9HEPC	Q68K42_9HEPC
65	13	11.0	659	2 Q68K41_9HEPC	Q68K41_9HEPC
66	13	11.0	659	2 Q68K40_9HEPC	Q68K40_9HEPC
67	13	11.0	659	2 Q68K39_9HEPC	Q68K39_9HEPC
68	13	11.0	659	2 Q68K37_9HEPC	Q68K37_9HEPC
69	13	11.0	659	2 Q68K35_9HEPC	Q68K35_9HEPC
70	13	11.0	659	2 Q68K33_9HEPC	Q68K33_9HEPC
71	13	11.0	659	2 Q68K32_9HEPC	Q68K32_9HEPC
72	13	11.0	659	2 Q68K31_9HEPC	Q68K31_9HEPC
73	13	11.0	659	2 Q68K30_9HEPC	Q68K30_9HEPC
74	13	11.0	660	2 Q68K29_9HEPC	Q68K29_9HEPC
75	13	11.0	660	2 Q68K28_9HEPC	Q68K28_9HEPC
76	13	11.0	661	2 Q68K27_9HEPC	Q68K27_9HEPC
77	13	11.0	661	2 Q68K26_9HEPC	Q68K26_9HEPC
78	13	11.0	2436	2 Q68K25_9HEPC	Q68K25_9HEPC
79	13	11.0	2742	2 Q68K24_9HEPC	Q68K24_9HEPC
80	13	11.0	2742	2 Q68K23_9HEPC	Q68K23_9HEPC
81	13	11.0	2908	2 Q68K22_9HEPC	Q68K22_9HEPC
82	13	11.0	3010	1 POLG_HCV1	POLG_HCV1
83	13	11.0	3010	1 POLG_HCVH	POLG_HCVH
84	13	11.0	3011	2 Q68K21_9HEPC	Q68K21_9HEPC
85	13	11.0	3011	2 Q68K20_9HEPC	Q68K20_9HEPC
86	13	11.0	3011	2 Q68K19_9HEPC	Q68K19_9HEPC
87	13	11.0	3011	2 Q68K18_9HEPC	Q68K18_9HEPC
88	13	11.0	3011	2 Q68K17_9HEPC	Q68K17_9HEPC
89	13	11.0	3011	2 Q68K16_9HEPC	Q68K16_9HEPC
90	13	11.0	3015	2 Q68K15_9HEPC	Q68K15_9HEPC
91	13	11.0	3015	2 Q68K14_9HEPC	Q68K14_9HEPC
92	12	10.2	138	2 Q68K13_9HEPC	Q68K13_9HEPC
93	12	10.2	138	2 Q68K12_9HEPC	Q68K12_9HEPC
94	12	10.2	138	2 Q68K11_9HEPC	Q68K11_9HEPC
95	12	10.2	138	2 Q68K10_9HEPC	Q68K10_9HEPC
96	12	10.2	138	2 Q68K09_9HEPC	Q68K09_9HEPC
97	12	10.2	138	2 Q68K08_9HEPC	Q68K08_9HEPC
98	12	10.2	138	2 Q68K07_9HEPC	Q68K07_9HEPC
99	12	10.2	138	2 Q68K06_9HEPC	Q68K06_9HEPC
100	12	10.2	138	2 Q68K05_9HEPC	Q68K05_9HEPC
101	12	10.2	138	2 Q68K04_9HEPC	Q68K04_9HEPC
102	12	10.2	138	2 Q68K03_9HEPC	Q68K03_9HEPC
103	12	10.2	138	2 Q68K02_9HEPC	Q68K02_9HEPC
104	12	10.2	138	2 Q68K01_9HEPC	Q68K01_9HEPC

105	12	10.2	138	2	Q68218	9HEPC	Q68218	hepatitis	C	178	12	10.2	3010	2	P90191	9HEPC	P90191	hepatitis	C
106	12	10.2	138	2	Q68221	9HEPC	Q68221	hepatitis	C	179	12	10.2	3010	2	P90193	9HEPC	P90193	hepatitis	C
107	12	10.2	138	2	Q68227	9HEPC	Q68227	hepatitis	C	180	12	10.2	3010	2	P90194	9HEPC	P90194	hepatitis	C
108	12	10.2	138	2	Q68228	9HEPC	Q68228	hepatitis	C	181	12	10.2	3010	2	Q02828	9HEPC	Q02828	hepatitis	C
109	12	10.2	138	2	Q68229	9HEPC	Q68229	hepatitis	C	182	12	10.2	3010	2	Q02829	9HEPC	Q02829	hepatitis	C
110	12	10.2	138	2	Q68230	9HEPC	Q68230	hepatitis	C	183	12	10.2	3010	2	Q5R2A8	9HEPC	Q5R2A8	hepatitis	C
111	12	10.2	138	2	Q68231	9HEPC	Q68231	hepatitis	C	184	12	10.2	3010	2	Q5R2A9	9HEPC	Q5R2A9	hepatitis	C
112	12	10.2	138	2	Q68232	9HEPC	Q68232	hepatitis	C	185	12	10.2	3010	2	Q5R2B0	9HEPC	Q5R2B0	hepatitis	C
113	12	10.2	138	2	Q68235	9HEPC	Q68235	hepatitis	C	186	12	10.2	3010	2	Q5R2B1	9HEPC	Q5R2B1	hepatitis	C
114	12	10.2	138	2	Q68237	9HEPC	Q68237	hepatitis	C	187	12	10.2	3010	2	Q5R2B2	9HEPC	Q5R2B2	hepatitis	C
115	12	10.2	138	2	Q68238	9HEPC	Q68238	hepatitis	C	188	12	10.2	3010	2	Q5R2B3	9HEPC	Q5R2B3	hepatitis	C
116	12	10.2	138	2	Q68240	9HEPC	Q68240	hepatitis	C	189	12	10.2	3010	2	Q5R2B4	9HEPC	Q5R2B4	hepatitis	C
117	12	10.2	138	2	Q68242	9HEPC	Q68242	hepatitis	C	190	12	10.2	3010	2	Q5R2B5	9HEPC	Q5R2B5	hepatitis	C
118	12	10.2	138	2	Q68244	9HEPC	Q68244	hepatitis	C	191	12	10.2	3010	2	Q5R2B6	9HEPC	Q5R2B6	hepatitis	C
119	12	10.2	172	2	Q81574	9HEPC	Q81574	hepatitis	C	192	12	10.2	3010	2	Q5R2B7	9HEPC	Q5R2B7	hepatitis	C
120	12	10.2	172	2	Q81575	9HEPC	Q81575	hepatitis	C	193	12	10.2	3010	2	Q5R2B8	9HEPC	Q5R2B8	hepatitis	C
121	12	10.2	172	2	Q81577	9HEPC	Q81577	hepatitis	C	194	12	10.2	3010	2	Q5R2B9	9HEPC	Q5R2B9	hepatitis	C
122	12	10.2	172	2	Q81578	9HEPC	Q81578	hepatitis	C	195	12	10.2	3010	2	Q5R2C0	9HEPC	Q5R2C0	hepatitis	C
123	12	10.2	172	2	Q81579	9HEPC	Q81579	hepatitis	C	196	12	10.2	3010	2	Q5R2C1	9HEPC	Q5R2C1	hepatitis	C
124	12	10.2	172	2	Q81580	9HEPC	Q81580	hepatitis	C	197	12	10.2	3010	2	Q5R2C2	9HEPC	Q5R2C2	hepatitis	C
125	12	10.2	172	2	Q81581	9HEPC	Q81581	hepatitis	C	198	12	10.2	3010	2	Q5R2C3	9HEPC	Q5R2C3	hepatitis	C
126	12	10.2	172	2	Q81582	9HEPC	Q81582	hepatitis	C	199	12	10.2	3010	2	Q5R2C4	9HEPC	Q5R2C4	hepatitis	C
127	12	10.2	172	2	Q81583	9HEPC	Q81583	hepatitis	C	200	12	10.2	3010	2	Q5R2C5	9HEPC	Q5R2C5	hepatitis	C
128	12	10.2	172	2	Q81584	9HEPC	Q81584	hepatitis	C	201	12	10.2	3010	2	Q5R2C6	9HEPC	Q5R2C6	hepatitis	C
129	12	10.2	271	2	Q81573	9HEPC	Q81573	hepatitis	C	202	12	10.2	3010	2	Q5R2C7	9HEPC	Q5R2C7	hepatitis	C
130	12	10.2	652	2	Q68K53	9HEPC	Q68K53	hepatitis	C	203	12	10.2	3010	2	Q5R2C8	9HEPC	Q5R2C8	hepatitis	C
131	12	10.2	659	2	Q68K59	9HEPC	Q68K59	hepatitis	C	204	12	10.2	3010	2	Q5R2C9	9HEPC	Q5R2C9	hepatitis	C
132	12	10.2	659	2	Q68K57	9HEPC	Q68K57	hepatitis	C	205	12	10.2	3010	2	Q5R2D0	9HEPC	Q5R2D0	hepatitis	C
133	12	10.2	659	2	Q68K50	9HEPC	Q68K50	hepatitis	C	206	12	10.2	3010	2	Q5R2D1	9HEPC	Q5R2D1	hepatitis	C
134	12	10.2	659	2	Q68K48	9HEPC	Q68K48	hepatitis	C	207	12	10.2	3010	2	Q5R2D2	9HEPC	Q5R2D2	hepatitis	C
135	12	10.2	659	2	Q68K44	9HEPC	Q68K44	hepatitis	C	208	12	10.2	3010	2	Q5R2D3	9HEPC	Q5R2D3	hepatitis	C
136	12	10.2	659	2	Q68K36	9HEPC	Q68K36	hepatitis	C	209	12	10.2	3010	2	Q5R2D4	9HEPC	Q5R2D4	hepatitis	C
137	12	10.2	660	2	Q68K46	9HEPC	Q68K46	hepatitis	C	210	12	10.2	3010	2	Q5R2D5	9HEPC	Q5R2D5	hepatitis	C
138	12	10.2	661	2	Q68K58	9HEPC	Q68K58	hepatitis	C	211	12	10.2	3010	2	Q5R2D6	9HEPC	Q5R2D6	hepatitis	C
139	12	10.2	661	2	Q68K52	9HEPC	Q68K52	hepatitis	C	212	12	10.2	3010	2	Q5R2D7	9HEPC	Q5R2D7	hepatitis	C
140	12	10.2	1805	2	Q41809	9HEPC	Q41809	hepatitis	C	213	12	10.2	3010	2	Q68788	9HEPC	Q68788	hepatitis	C
141	12	10.2	1984	2	Q774V8	9HEPC	Q774V8	hepatitis	C	214	12	10.2	3010	2	Q68826	9HEPC	Q68826	hepatitis	C
142	12	10.2	2284	2	Q81817	9HEPC	Q81817	hepatitis	C	215	12	10.2	3010	2	Q68833	9HEPC	Q68833	hepatitis	C
143	12	10.2	2864	2	Q92973	9HEPC	Q92973	hepatitis	C	216	12	10.2	3010	2	Q68949	9HEPC	Q68949	hepatitis	C
144	12	10.2	2864	2	Q92974	9HEPC	Q92974	hepatitis	C	217	12	10.2	3010	2	Q81541	9HEPC	Q81541	hepatitis	C
145	12	10.2	2864	2	Q92975	9HEPC	Q92975	hepatitis	C	218	12	10.2	3010	2	Q81757	9HEPC	Q81757	hepatitis	C
146	12	10.2	2864	2	Q92976	9HEPC	Q92976	hepatitis	C	219	12	10.2	3010	2	Q81760	9HEPC	Q81760	hepatitis	C
147	12	10.2	2864	2	Q9WLK8	9HEPC	Q9WLK8	hepatitis	C	220	12	10.2	3010	2	Q81989	9HEPC	Q81989	hepatitis	C
148	12	10.2	2864	2	Q9WLK9	9HEPC	Q9WLK9	hepatitis	C	221	12	10.2	3010	2	Q8V638	9HEPC	Q8V638	hepatitis	C
149	12	10.2	2864	2	Q9WLLO	9HEPC	Q9WLLO	hepatitis	C	222	12	10.2	3010	2	Q99AU2	9HEPC	Q99AU2	hepatitis	C
150	12	10.2	3008	2	Q9J3F4	9HEPC	Q9J3F4	hepatitis	C	223	12	10.2	3010	2	Q9DTD6	9HEPC	Q9DTD6	hepatitis	C
151	12	10.2	3009	1	POLG_HCVBK		P25663	h genome	po	224	12	10.2	3010	2	Q9DTD7	9HEPC	Q9DTD7	hepatitis	C
152	12	10.2	3009	1	POLG_HCVCO		Q9Vmx2	h genome	po	225	12	10.2	3010	2	Q9DTD9	9HEPC	Q9DTD9	hepatitis	C
153	12	10.2	3009	1	POLG_HCVJA		P25662	h genome	po	226	12	10.2	3010	2	Q9DTE1	9HEPC	Q9DTE1	hepatitis	C
154	12	10.2	3009	1	POLG_HCVUT		Q00269	h genome	po	227	12	10.2	3010	2	Q9DTE2	9HEPC	Q9DTE2	hepatitis	C
155	12	10.2	3009	1	POLG_HCVTW		P29846	h genome	po	228	12	10.2	3010	2	Q9DTE3	9HEPC	Q9DTE3	hepatitis	C
156	12	10.2	3010	2	Q59IP0	9HEPC	Q59IP0	hepatitis	C	229	12	10.2	3010	2	Q9DTE4	9HEPC	Q9DTE4	hepatitis	C
157	12	10.2	3010	2	Q68285	9HEPC	Q68285	hepatitis	C	230	12	10.2	3010	2	Q9DTE5	9HEPC	Q9DTE5	hepatitis	C
158	12	10.2	3010	2	Q68533	9HEPC	Q68533	hepatitis	C	231	12	10.2	3010	2	Q9DTE7	9HEPC	Q9DTE7	hepatitis	C
159	12	10.2	3010	2	Q6GYR9	9HEPC	Q6GYR9	hepatitis	C	232	12	10.2	3010	2	Q9DTE9	9HEPC	Q9DTE9	hepatitis	C
160	12	10.2	3010	2	Q807P3	9HEPC	Q807P3	hepatitis	C	233	12	10.2	3010	2	Q9DTF0	9HEPC	Q9DTF0	hepatitis	C
161	12	10.2	3010	2	Q81825	9HEPC	Q81825	hepatitis	C	234	12	10.2	3010	2	Q9J3F9	9HEPC	Q9J3F9	hepatitis	C
162	12	10.2	3010	2	Q8QRL8	9HEPC	Q8QRL8	hepatitis	C	235	12	10.2	3010	2	Q9J3G0	9HEPC	Q9J3G0	hepatitis	C
163	12	10.2	3010	2	Q913V3	9HEPC	Q913V3	hepatitis	C	236	12	10.2	3010	2	Q9J3G1	9HEPC	Q9J3G1	hepatitis	C
164	12	10.2	3010	2	Q9DTE6	9HEPC	Q9DTE6	hepatitis	C	237	12	10.2	3010	2	Q9J3G2	9HEPC	Q9J3G2	hepatitis	C
165	12	10.2	3010	2	Q9J3G8	9HEPC	Q9J3G8	hepatitis	C	238	12	10.2	3010	2	Q9J3G3	9HEPC	Q9J3G3	hepatitis	C
166	12	10.2	3010	2	Q9J3H1	9HEPC	Q9J3H1	hepatitis	C	239	12	10.2	3010	2	Q9J3G4	9HEPC	Q9J3G4	hepatitis	C
167	12	10.2	3010	2	Q9J3H3	9HEPC	Q9J3H3	hepatitis	C	240	12	10.2	3010	2	Q9J3G5	9HEPC	Q9J3G5	hepatitis	C
168	12	10.2	3010	2	Q9QIX3	9HEPC	Q9QIX3	hepatitis	C	241	12	10.2	3010	2	Q9J3G6	9HEPC	Q9J3G6	hepatitis	C
169	12	10.2	3010	2	P88803	9HEPC	P88803	hepatitis	C	242	12	10.2	3010	2	Q9J3G7	9HEPC	Q9J3G7	hepatitis	C
170	12	10.2	3010	2	Q4PKP9	9HEPC	Q4PKP9	hepatitis	C	243	12	10.2	3010	2	Q9J3G9	9HEPC	Q9J3G9	hepatitis	C
171	12	10.2	3010	2	Q9J796	9HEPC	Q9J796	hepatitis	C	244	12	10.2	3010	2	Q9J3H0	9HEPC	Q9J3H0	hepatitis	C
172	12	10.2	3010	2	Q92969	9HEPC	Q92969	hepatitis	C	245	12	10.2	3010	2	Q9J3H2	9HEPC	Q9J3H2	hepatitis	C
173	12	10.2	3010	2	Q92970	9HEPC	Q92970	hepatitis	C	246	12	10.2	3010	2	Q9J3H5	9HEPC	Q9J3H5	hepatitis	C
174	12	10.2	3010	2	Q92971	9HEPC	Q92971	hepatitis	C	247	12	10.2	3010	2	Q9J3H6	9HEPC	Q9J3H6	hepatitis	C
175	12	10.2	3010	2	Q92972	9HEPC	Q92972	hepatitis	C	248	12	10.2	3010	2	Q9J3H7	9HEPC	Q9J3H7	hepatitis	C
176	12	10.2	3010	2	Q93016	9HEPC	Q93016	hepatitis	C	249	12	10.2	3010	2	Q9J3H8	9HEPC	Q9J3H8	hepatitis	C
177	12	10.2	3010	2	Q93077	9HEPC	Q93077	hepatitis	C	250	12	10.2	3010	2	Q9J3H9	9HEPC	Q9J3H9	hepatitis	C

251	12	10.2	3010	2	Q9J310_9HEPC	Q9J310	hepatitis c	324	8	6.8	560	2	Q4H6D4_9DEIO	Q4H6d4	deinococcus
252	12	10.2	3010	2	Q9J311_9HEPC	Q9J311	hepatitis c	325	8	6.8	592	2	Q8EH53_SHEON	Q8eh53	shewanella
253	12	10.2	3010	2	Q9Q1X1_9HEPC	Q9Q1x1	hepatitis c	326	8	6.8	631	2	Q4JME8_9HEPC	Q4jme8	hepatitis c
254	12	10.2	3010	2	Q9Q1X2_9HEPC	Q9Q1x2	hepatitis c	327	8	6.8	631	2	Q4JME9_9HEPC	Q4jme9	hepatitis c
255	12	10.2	3010	2	Q9Q1X4_9HEPC	Q9Q1x4	hepatitis c	328	8	6.8	631	2	Q4JMF0_9HEPC	Q4jmf0	hepatitis c
256	12	10.2	3010	2	Q9Q1X5_9HEPC	Q9Q1x5	hepatitis c	329	8	6.8	631	2	Q4JMF1_9HEPC	Q4jmf1	hepatitis c
257	12	10.2	3010	2	Q9Q1X6_9HEPC	Q9Q1x6	hepatitis c	330	8	6.8	631	2	Q4JMF2_9HEPC	Q4jmf2	hepatitis c
258	12	10.2	3010	2	Q9Q1X7_9HEPC	Q9Q1x7	hepatitis c	331	8	6.8	631	2	Q4JMF3_9HEPC	Q4jmf3	hepatitis c
259	12	10.2	3010	2	Q9Q1X8_9HEPC	Q9Q1x8	hepatitis c	332	8	6.8	631	2	Q4JMF4_9HEPC	Q4jmf4	hepatitis c
260	12	10.2	3010	2	Q9Q1Y1_9HEPC	Q9Q1y1	hepatitis c	333	8	6.8	631	2	Q4JMF5_9HEPC	Q4jmf5	hepatitis c
261	12	10.2	3010	2	Q9Q1Y2_9HEPC	Q9Q1y2	hepatitis c	334	8	6.8	631	2	Q4JMF6_9HEPC	Q4jmf6	hepatitis c
262	12	10.2	3010	2	Q9Q1Y3_9HEPC	Q9Q1y3	hepatitis c	335	8	6.8	631	2	Q4JMF7_9HEPC	Q4jmf7	hepatitis c
263	12	10.2	3010	2	Q9Q1Y4_9HEPC	Q9Q1y4	hepatitis c	336	8	6.8	631	2	Q4JMF8_9HEPC	Q4jmf8	hepatitis c
264	12	10.2	3010	2	Q9Q1Y5_9HEPC	Q9Q1y5	hepatitis c	337	8	6.8	631	2	Q4JMF9_9HEPC	Q4jmf9	hepatitis c
265	12	10.2	3010	2	Q9Q1Y6_9HEPC	Q9Q1y6	hepatitis c	338	8	6.8	631	2	Q4JMG0_9HEPC	Q4jmg0	hepatitis c
266	12	10.2	3010	2	Q9Q1Y7_9HEPC	Q9Q1y7	hepatitis c	339	8	6.8	631	2	Q4JMG1_9HEPC	Q4jmg1	hepatitis c
267	12	10.2	3010	2	Q9Q1Y8_9HEPC	Q9Q1y8	hepatitis c	340	8	6.8	631	2	Q4JMG2_9HEPC	Q4jmg2	hepatitis c
268	12	10.2	3010	2	Q9Q1Y9_9HEPC	Q9Q1y9	hepatitis c	341	8	6.8	631	2	Q4JMG3_9HEPC	Q4jmg3	hepatitis c
269	12	10.2	3010	2	Q9Q1Z0_9HEPC	Q9Q1z0	hepatitis c	342	8	6.8	631	2	Q4JMG4_9HEPC	Q4jmg4	hepatitis c
270	12	10.2	3010	2	Q9QP06_9HEPC	Q9QP06	hepatitis c	343	8	6.8	631	2	Q4JMG5_9HEPC	Q4jmg5	hepatitis c
271	12	10.2	3010	2	Q9QP61_9HEPC	Q9QP61	hepatitis c	344	8	6.8	631	2	Q4JMG6_9HEPC	Q4jmg6	hepatitis c
272	12	10.2	3010	2	Q9WIK8_9HEPC	Q9WIK8	hepatitis c	345	8	6.8	631	2	Q4JMG7_9HEPC	Q4jmg7	hepatitis c
273	12	10.2	3011	2	Q6SCJ5_9HEPC	Q6scj5	hepatitis c	346	8	6.8	631	2	Q4JMG8_9HEPC	Q4jmg8	hepatitis c
274	12	10.2	3011	2	Q9DTE8_9HEPC	Q9dte8	hepatitis c	347	8	6.8	631	2	Q4JMG9_9HEPC	Q4jmg9	hepatitis c
275	12	10.2	3011	2	Q9DTE3_9HEPC	Q9dte3	hepatitis c	348	8	6.8	631	2	Q4JMH0_9HEPC	Q4jmh0	hepatitis c
276	12	10.2	3012	2	Q9WIK7_9HEPC	Q9WIK7	hepatitis c	349	8	6.8	631	2	Q4JMH1_9HEPC	Q4jmh1	hepatitis c
277	12	10.2	3012	2	Q6J6P5_9HEPC	Q6j6p5	hepatitis c	350	8	6.8	631	2	Q4JMH2_9HEPC	Q4jmh2	hepatitis c
278	12	10.2	3013	2	Q9J3H4_9HEPC	Q9J3h4	hepatitis c	351	8	6.8	631	2	Q4JMH3_9HEPC	Q4jmh3	hepatitis c
279	12	10.2	3013	2	Q9Q1X9_9HEPC	Q9Q1x9	hepatitis c	352	8	6.8	631	2	Q4JMH4_9HEPC	Q4jmh4	hepatitis c
280	12	10.2	3013	2	Q9Q1Y0_9HEPC	Q9Q1y0	hepatitis c	353	8	6.8	631	2	Q4JMH5_9HEPC	Q4jmh5	hepatitis c
281	12	10.2	3013	2	Q9QNC0_9HEPC	Q9QNC0	hepatitis c	354	8	6.8	631	2	Q4JMH6_9HEPC	Q4jmh6	hepatitis c
282	12	10.2	3014	2	Q6GYR8_9HEPC	Q6GYR8	hepatitis c	355	8	6.8	631	2	Q4JMH7_9HEPC	Q4jmh7	hepatitis c
283	12	10.2	3014	2	Q86614_9HEPC	Q86614	hepatitis c	356	8	6.8	657	2	Q4ZRO2_PSES	Q4zr02	pseudomonas
284	12	10.2	3014	2	Q9DTE0_9HEPC	Q9dte0	hepatitis c	357	8	6.8	657	2	Q87Z08_PSESM	Q87z08	pseudomonas
285	12	10.2	3015	2	Q9WPH5_9HEPC	Q9wph5	hepatitis c	358	8	6.8	707	2	Q4PH79_USTMA	Q4ph79	usilago ma
286	12	10.2	3019	2	Q92529_9HEPC	Q92529	hepatitis c	359	8	6.8	713	2	Q5YPX2_NOCFA	Q5ypx2	nocardia fa
287	11	9.3	3010	2	Q91AU0_9HEPC	Q91AU0	hepatitis c	360	8	6.8	760	2	Q5AYN4_EMENI	Q5ayn4	aspergillus
288	11	9.3	3022	2	Q68798_9HEPC	Q68798	hepatitis c	361	8	6.8	3010	2	P90192_9HEPC	P90192	hepatitis c
289	10	8.5	434	2	Q988U6_RHILO	Q988u6	rhizobium l	362	8	6.8	3010	2	P90195_9HEPC	P90195	hepatitis c
290	9	7.6	280	2	Q9HRP0_HALSA	Q9hrp0	halobacteri	363	8	6.8	3013	2	Q92530_9HEPC	Q92530	hepatitis c
291	9	7.6	484	2	Q81242_9HEPC	Q81242	hepatitis c	364	7	5.9	61	2	Q51DU1_ENTHI	Q51du1	entamoeba h
292	9	7.6	545	2	Q521Y4_NOCFA	Q521y4	nocardia fa	365	7	5.9	71	2	Q98JEL_RHILO	Q98jel	rhizobium l
293	9	7.6	1235	2	Q6AH48_LEIXX	Q6ah48	leifsonia x	366	7	5.9	76	2	Q9AF06_9ACTO	Q9af06	frankia sp.
294	9	7.6	3014	2	Q39928_9HEPC	Q39928	hepatitis c	367	7	5.9	82	1	GCHI_BUGGR	Gchi	bugger
295	9	7.6	3014	2	Q91936_9HEPC	Q91936	hepatitis c	368	7	5.9	82	2	Q8VJS5_MYCTU	Q8vjs5	mycobacteri
296	8	6.8	117	2	Q4H710_9DEIO	Q4h710	deinococcus	369	7	5.9	82	2	Q9Q3S2_9HEPC	Q9q3s2	hepatitis c
297	8	6.8	138	2	Q68234_9HEPC	Q68234	hepatitis c	370	7	5.9	87	2	Q7NH49_GLOVI	Q7nh49	gloeobacter
298	8	6.8	186	2	Q4P348_USTMA	Q4p348	usilago ma	371	7	5.9	90	2	Q39915_9HEPC	Q39915	hepatitis c
299	8	6.8	215	2	Q4UXM7_XANCP	Q4uxm7	xanthomonas	372	7	5.9	91	2	Q5SK97_THET8	Q5sk97	thermus the
300	8	6.8	215	2	Q8P6D8_XANCP	Q8p6d8	xanthomonas	373	7	5.9	92	2	Q532L0_9HEPC	Q532l0	hepatitis c
301	8	6.8	220	2	Q4NV17_9DELT	Q4nv17	anaeromyxob	374	7	5.9	96	2	Q9HVF0_PSEAE	Q9hvf0	pseudomonas
302	8	6.8	279	2	Q93HX1_MAGMG	Q93hx1	magnetospir	375	7	5.9	102	2	Q6ACG9_LEIXX	Q6acg9	leifsonia x
303	8	6.8	296	2	Q8PWS7_METMA	Q8pws7	methanosaer	376	7	5.9	102	2	Q9PXP4_9HEPC	Q9pxp4	hepatitis c
304	8	6.8	305	2	Q5YXT2_NOCFA	Q5yxt2	nocardia fa	377	7	5.9	103	2	Q61GW9_CAEBR	Q61gw9	caenorhabdi
305	8	6.8	320	2	Q9L1Y9_STRCO	Q9l1y9	streptomyce	378	7	5.9	104	2	Q6N7U3_RHOPA	Q6n7u3	rhodopseudo
306	8	6.8	327	1	PHO36_YEAST	Q12442	saccharomyc	379	7	5.9	107	2	Q915D6_PSEAE	Q915d6	pseudomonas
307	8	6.8	342	2	Q57F22_BRUAB	Q57f22	brucella ab	380	7	5.9	107	2	Q5B0J4_EMENI	Q5b0j4	aspergillus
308	8	6.8	342	2	Q8G2I3_BRUSU	Q8g2i3	brucella su	381	7	5.9	117	2	Q5B0J4_EMENI	Q5b0j4	aspergillus
309	8	6.8	342	2	Q8YFD6_BRUME	Q8yfd6	brucella me	382	7	5.9	121	2	Q9SC1_BRAJA	Q9sc1	bradyrhizob
310	8	6.8	363	2	Q31809_BACSU	Q31809	bacillus su	383	7	5.9	122	2	Q7NKD5_GLOVI	Q7nkd5	gloeobacter
311	8	6.8	369	2	Q82CF6_STRAW	Q82cf6	streptomyce	384	7	5.9	125	2	Q7NEL9_GLOVI	Q7nel9	gloeobacter
312	8	6.8	399	2	Q7DAN4_MYCTU	Q7dan4	mycobacteri	385	7	5.9	125	2	Q81592_9HEPC	Q81592	hepatitis c
313	8	6.8	399	2	Q7TVB9_MYCBO	Q7tve9	mycobacteri	386	7	5.9	128	1	NUOA_MYCTU	NUOA	mycobacteri
314	8	6.8	399	2	Q79FP0_MYCTU	Q79fp0	mycobacteri	387	7	5.9	128	2	Q956X0_9PLAT	Q956x0	choricotype
315	8	6.8	401	2	Q8J7J0_VTBPA	Q8j7j0	vibrio para	388	7	5.9	128	2	Q956X0_9PLAT	Q956x0	choricotype
316	8	6.8	419	1	CBPAL_FIG	P09954	su scrofa	389	7	5.9	128	2	Q8NRG8_9DELT	Q8nrg8	anaeromyxob
317	8	6.8	420	2	Q9HS33_HALSA	Q9hs33	halobacteri	390	7	5.9	128	2	Q89SL2_BRAJA	Q89sl2	bradyrhizob
318	8	6.8	425	2	Q68344_9HEPC	Q68344	hepatitis c	391	7	5.9	131	2	Q63P16_BURPS	Q63p16	burkholderi
319	8	6.8	449	2	Q9QP2_9HEPC	Q9qp2	hepatitis c	392	7	5.9	132	2	Q956I8_9BASI	Q956i8	rhodotorula
320	8	6.8	449	2	Q9WNM3_9HEPC	Q9wnm3	hepatitis c	393	7	5.9	132	2	Q956H5_9BASI	Q956h5	rhodospo
321	8	6.8	449	2	Q9QCP1_9HEPC	Q9qcp1	hepatitis c	394	7	5.9	132	2	Q956I6_9BASI	Q956i6	rhodotorula
322	8	6.8	449	2	Q9QCP3_9HEPC	Q9qcp3	hepatitis c	395	7	5.9	132	2	Q956K6_9BASI	Q956k6	rhodotorula
323	8	6.8	559	2	Q4NTJ3_9DELT	Q4ntj3	anaeromyxob	396	7	5.9	138	2	Q68214_9HEPC	Q68214	hepatitis c

397	7	5.9	138	2	Q68219_9HEPC	Q68219	hepatitis c
398	7	5.9	138	2	Q68220_9HEPC	Q68220	hepatitis c
399	7	5.9	142	2	Q5J6J6_MYCVN	Q5J6J6	mycobacteri
400	7	5.9	146	2	Q99KF4_MOUSE	Q99KF4	mus musculus
401	7	5.9	147	2	Q8TW80_METKA	Q8TW80	magnaporthe
402	7	5.9	151	2	Q521A9_NAGGR	Q521A9	magnaporthe
403	7	5.9	155	2	Q72BY0_DESVH	Q72BY0	desulfovibr
404	7	5.9	159	2	Q57VK1_9TRYP	Q57VK1	trypanosoma
405	7	5.9	159	2	Q92WY2_RHIME	Q92WY2	rhizobium m
406	7	5.9	160	2	Q7YGT5_9GOBI	Q7YGT5	gnatholepis
407	7	5.9	161	2	Q4NR09_9DELT	Q4NR09	anaeromyxob
408	7	5.9	168	2	Q514C4_PECGU	Q514C4	pectinaria
409	7	5.9	172	2	Q9HQ7_9HALSA	Q9HQ7	halobacteri
410	7	5.9	174	1	Y433_AERPE	Y433	aeropyrum p
411	7	5.9	177	2	Q5P936_AZOSE	Q5P936	azocarcus sp
412	7	5.9	178	2	Q67LJ9_SYMTH	Q67LJ9	syntrophobacte
413	7	5.9	179	1	FAIMI_MOUSE	FAIMI	mouse
414	7	5.9	179	2	Q8ECD8_SHEON	Q8ECD8	shewanella
415	7	5.9	184	2	Q84X29_BRARP	Q84X29	brassica ra
416	7	5.9	184	2	Q73VJ0_MYCPA	Q73VJ0	mycobacteri
417	7	5.9	185	2	Q7U864_SYNXP	Q7U864	synecococc
418	7	5.9	185	2	Q68QF2_NVPAP	Q68QF2	anthraerea p
419	7	5.9	186	2	Q91XV9_9HEPC	Q91XV9	hepatitis c
420	7	5.9	187	2	Q8ORV4_9BETA	Q8ORV4	pongine her
421	7	5.9	188	2	Q64D23_9ARCH	Q64D23	uncultured
422	7	5.9	188	2	Q57D47_BRUAB	Q57D47	brucella ab
423	7	5.9	188	2	Q8GOK2_BRUSU	Q8GOK2	brucella su
424	7	5.9	188	2	Q9YH81_BRUME	Q9YH81	brucella me
425	7	5.9	190	1	GC11_PSEPK	GC11	pseudomonas
426	7	5.9	192	1	Q5QMN8_ORYSA	Q5QMN8	oryza sativ
427	7	5.9	192	2	Q9LCT1_BRAJA	Q9LCT1	bradyrhizob
428	7	5.9	197	2	Q7NPH4_GLOVI	Q7NPH4	gloeobacter
429	7	5.9	203	1	VP10_BPPRD	VP10	bacterioph
430	7	5.9	203	2	Q6EDX4_BPPRD	Q6EDX4	bacterioph
431	7	5.9	204	2	Q6N9H6_RHOPA	Q6N9H6	rhodosphe
432	7	5.9	204	2	Q930S6_RHIME	Q930S6	rhizobium m
433	7	5.9	206	2	Q8BZ84_SCHJA	Q8BZ84	schistosoma
434	7	5.9	215	2	Q4WFS9_ASPFU	Q4WFS9	aspergillus
435	7	5.9	217	2	Q9J7E1_NEIG1	Q9J7E1	neisseria g
436	7	5.9	217	2	Q9JTIO_NEIMA	Q9JTIO	neisseria m
437	7	5.9	217	2	Q9JYH9_NEIMB	Q9JYH9	neisseria m
438	7	5.9	219	2	Q9V4H2_DROME	Q9V4H2	drosophila
439	7	5.9	219	2	Q4ZVP9_PSESY	Q4ZVP9	pseudomonas
440	7	5.9	222	2	Q8DNPT_STRK6	Q8DNPT	streptococc
441	7	5.9	222	2	Q97P66_STKPN	Q97P66	streptococc
442	7	5.9	224	2	Q8NS26_CORGL	Q8NS26	corynebacte
443	7	5.9	224	2	Q5TBB2_HUMAN	Q5TBB2	homo sapien
444	7	5.9	224	2	Q6N4M2_RHOPA	Q6N4M2	rhodosphe
445	7	5.9	226	2	Q8A6N6_BACTN	Q8A6N6	bacteroides
446	7	5.9	228	1	ISPD_BACHD	ISPD	bacillus ha
447	7	5.9	231	2	Q8A1G6_BACTN	Q8A1G6	bacteroides
448	7	5.9	234	2	Q97MD2_CLOAB	Q97MD2	clostridium
449	7	5.9	235	1	YOGA_ECOLI	YOGA	escherichia
450	7	5.9	235	2	Q83Q81_SHIFL	Q83Q81	shigella fl
451	7	5.9	236	2	Q8V817_KLEOX	Q8V817	klebsiella
452	7	5.9	245	2	Q8YQ77_ANASP	Q8YQ77	anabaena sp
453	7	5.9	246	2	Q63QR8_BURPS	Q63QR8	burkholderi
454	7	5.9	247	2	Q5YXG7_NOCFA	Q5YXG7	nocardia fa
455	7	5.9	250	2	Q75CX5_ASHGO	Q75CX5	ashbya gos
456	7	5.9	250	2	Q75DB2_ASHGO	Q75DB2	ashbya gos
457	7	5.9	253	1	ISPD_RALSO	ISPD	ralesonia s
458	7	5.9	253	2	Q4S101_TETNG	Q4S101	tetradodon n
459	7	5.9	255	2	Q62H02_BURMA	Q62H02	burkholderi
460	7	5.9	256	2	Q81586_9HEPC	Q81586	hepatitis c
461	7	5.9	259	1	HURE_METWA	HURE	methanosarc
462	7	5.9	259	2	Q41Q33_9BURK	Q41Q33	burkholderi
463	7	5.9	261	2	Q75DB1_ASHGO	Q75DB1	ashbya gos
464	7	5.9	262	1	THIM_STAEP	THIM	staphylococ
465	7	5.9	262	1	THIM_STAEP	THIM	staphylococ
466	7	5.9	264	2	Q4NSY4_9BOKA	Q4NSY4	geobacillus
467	7	5.9	267	2	Q4NSY4_9BELT	Q4NSY4	anaeromyxob
468	7	5.9	272	2	Q4LJA5_9BURK	Q4LJA5	burkholderi
469	7	5.9	274	2	Q4NTJ8_9BELT	Q4NTJ8	anaeromyxob

Q7P114	chromobacte	2	Q7P114_CHRVO	2	275	7	5.9
P54993	streptomyce	1	SNAB_STRPR	1	276	7	5.9
Q8ZW5	pyrobaculum	2	Q8ZWH5_PYRAR	2	276	7	5.9
Q83W1	salmonella	2	Q83W1_SALT1	2	277	7	5.9
Q8ZLV5	salmonella	2	Q8ZLV5_SALPA	2	277	7	5.9
Q5P187	salmonella	2	Q5P187_SALPA	2	277	7	5.9
Q8GQ1	pseudomonas	2	Q8GOA1_PSEAE	2	278	7	5.9
Q8DA19	vibrio vuln	2	Q8DA19_VIBVU	2	279	7	5.9
Q7MJ12	vibrio vuln	2	Q7MJ12_VIBVU	2	279	7	5.9
Q83YD5	streptomyce	2	Q83YD5_STRHY	2	280	7	5.9
Q93J38	streptomyce	2	Q93J38_STRCO	2	284	7	5.9
Q6A624	propionibac	2	Q6A624_PROAC	2	284	7	5.9
Q8G2Y2	brucella su	2	Q8G2Y2_BRUSU	2	285	7	5.9
Q8PVM8	methanosarc	2	Q8PVM8_METWA	2	286	7	5.9
Q8RKH3	streptomyce	2	Q8RKH3_STRCO	2	288	7	5.9
Q6BMQ0	debaromyce	2	Q6BMQ0_DEBHA	2	288	7	5.9
Q6ACJ2	leifsonia x	2	Q6ACJ2_LEIXX	2	288	7	5.9
Q511G7	silicibacte	2	Q511G7_SILPO	2	288	7	5.9
Q67RA6	syntrophobacte	2	Q67RA6_SYMTH	2	289	7	5.9
Q89G69	bradyrhizob	2	Q89G69_BRAJA	2	293	7	5.9
P56218	pyrococcus	1	AMPM_PYRFU	1	295	7	5.9
O58362	pyrococcus	1	AMPM_PYRHO	1	295	7	5.9
Q5JGD1	pyrococcus	2	Q5JGD1_PYRKO	2	295	7	5.9
Q62KT6	burkholderi	2	Q62KT6_BURMA	2	297	7	5.9
Q5UXA2	haloarcula	2	Q5UXA2_HALMA	2	298	7	5.9
Q83UM4	burkholderi	2	Q83UM4_BURFS	2	300	7	5.9
Q7U8H1	synecococc	2	Q7U8H1_SYNFX	2	302	7	5.9
Q57FK0	brucella ab	2	Q57FK0_BRUAB	2	303	7	5.9
Q8YEV7	brucella me	2	Q8YEV7_BRUME	2	303	7	5.9
P65179	mycobacteri	1	ISPB_MYCBO	1	306	7	5.9
P65178	mycobacteri	1	ISPB_MYCTU	1	306	7	5.9
Q5YCE0	acinetobact	1	QTC_ACIAAD	1	306	7	5.9
Q8YCE0	aeropyrum p	2	Q9YCE0_AERPE	2	306	7	5.9
Q4HCY3	deinococcus	2	Q4HCY3_9DEIO	2	306	7	5.9
Q5GY13	xanthomonas	2	Q5GY13_XANOR	2	307	7	5.9
Q61Q36	caenorhabdi	2	Q61Q36_CAEBR	2	310	7	5.9
O5YWF0	nocardia fa	2	O5YWF0_NOCFA	2	310	7	5.9
O33744	streptomyce	2	O33744_STRSQ	2	312	7	5.9
Q4J105	azotobacter	2	Q4J105_AZOVI	2	312	7	5.9
Q4KI50	pseudomonas	2	Q4KI50_PSEF5	2	313	7	5.9
TRMB	PROAC	1	TRMB_PROAC	1	315	7	5.9
Q89L42	bradyrhizob	2	Q89L42_BRAJA	2	315	7	5.9
Q741W1	mycobacteri	1	ISPB_MYCPA	1	316	7	5.9
Q9VBN1	drosophila	2	Q9VBN1_DROME	2	316	7	5.9
Q5TBB3	homo sapien	2	Q5TBB3_HUMAN	2	319	7	5.9
Q7Z4Z9	homo sapien	2	Q7Z4Z9_HUMAN	2	319	7	5.9
Q5RDZ2	pongo pygma	2	Q5RDZ2_PONPY	2	319	7	5.9
Q727G8	desulfovibr	2	Q727G8_DESVH	2	322	7	5.9
Q72DA2	desulfovibr	2	Q72DA2_DESVH	2	323	7	5.9
Q4TN16	erythrobact	2	Q4TN16_9SPHN	2	327	7	5.9
Q89R48	bradyrhizob	2	Q89R48_BRAJA	2	328	7	5.9
Q5TY15	anophelis g	2	Q5TY15_ANOGA	2	331	7	5.9
Q82PF3	streptomyce	2	Q82PF3_STRAW	2	332	7	5.9
Q60GF1	comamonas a	2	Q60GF1_COMAC	2	335	7	5.9
Q6TFJ0	pseudomonas	2	Q6TFJ0_PSESD	2	335	7	5.9
Q9EXM3	pseudomonas	2	Q9EXM3_PSESP	2	335	7	5.9
Q55N25	cryptococcu	2	Q55N25_CRYNE	2	339	7	5.9
Q5KBF4	cryptococcu	2	Q5KBF4_CRYNE	2	339	7	5.9
Q72TU1	leptospira	2	Q72TU1_LEPIC	2	340	7	5.9
Q8F1D2	leptospira	2	Q8F1D2_LEPIN	2	340	7	5.9
Y0401	homo sapien	1	Y0401_HUMAN	1	345	7	5.9
Q7XT06	oryza sativ	2	Q7XT06_ORYSA	2	345	7	5.9
Q73X82	mycobacteri	2	Q73X82_MYCPA	2	347	7	5.9
PYRD	XYLFA	1	PYRD_XYLFA	1	351	7	5.9
PYRD	XYLFT	1	PYRD_XYLFT	1	351	7	5.9
Q60KH8	caenorhabdi	2	Q60KH8_CAEBR	2	354	7	5.9
Q82XV2	nitrosomona	2	Q82XV2_NITEU	2	354	7	5.9
Q7V810	prochloroco	2	Q7V810_PROMA	2	356	7	5.9
Q92T67	rhizobium m	2	Q92T67_RHIME	2	357	7	5.9
Q7WH91	borderetella	2	Q7WH91_BORPE	2	360	7	5.9
Q7VWM7	borderetella	2	Q7VWM7_BORPE	2	361	7	5.9
Q7W9H1	borderetella	2	Q7W9H1_BORPA	2	361	7	5.9
Q4ND64	arthrobacte	2	Q4ND64_9NICC	2	362	7	5.9

543	7	5.9	362	2	054158_STRCO	054158 streptomyc	616	7	5.9	435	2	07XTK3_ORYSA	Q7xtk3 oryza sativ
544	7	5.9	364	2	04HPY0_CAMCO	Q4hpy0 campylobact	617	7	5.9	435	2	073UN2_MYCPA	Q73un2 mycobacteri
545	7	5.9	366	1	Y1821_SYNY3	P73714 synchocyst	618	7	5.9	437	2	08IG53_CAREL	Q8ig53 caenorhabdi
546	7	5.9	367	2	0682F3_ARATH	Q682f3 arabidopsis	619	7	5.9	437	2	07NJ53_GLOVI	Q7nje3 gloeobacte
547	7	5.9	367	2	Q888S1_ARATH	Q888s1 arabidopsis	620	7	5.9	443	2	Q7NQS8_COREF	Q7nq8 corynebacte
548	7	5.9	367	2	Q9XEG0_ARATH	Q9xeg0 arabidopsis	621	7	5.9	449	2	Q5H4B3_XANOR	Q5h4b3 xanthomonas
549	7	5.9	370	2	Q9YCN9_AERPE	Q9ycn9 aeropyrum p	622	7	5.9	450	1	AROA_MYCBO	Q7twy4 mycobacteri
550	7	5.9	374	2	Q7U3W6_SYNPK	Q7u3w6 synchococc	623	7	5.9	450	1	AROA_MYCTU	P22a87 mycobacteri
551	7	5.9	375	2	Q4LX47_9BURK	Q4lx47 burkholderi	624	7	5.9	450	2	Q6O3J2_METCA	Q6o3j2 methylococc
552	7	5.9	375	2	Q7V4A0_PROHM	Q7v4a0 prochloroco	625	7	5.9	451	2	Q5E7W9_VIBF1	Q5e7w9 vibrio fisc
553	7	5.9	379	2	Q94X28_9CICH	Q94x28 rtychochrom	626	7	5.9	451	2	Q8DE98_VIBPU	Q8de98 vibrio vuln
554	7	5.9	382	2	Q6NAN4_RHOPA	Q6nan4 rhodopseudo	627	7	5.9	451	2	Q87LD7_VIBPA	Q87ld7 vibrio para
555	7	5.9	383	2	Q01739_CAREL	Q01739 caenorhabdi	628	7	5.9	451	2	Q7MPC2_VIBVY	Q7mpc2 vibrio para
556	7	5.9	383	2	Q89DB5_BRAJA	Q89db5 bradyrhizob	629	7	5.9	452	2	Q6LMB1_PHOPR	Q6lmb1 photobacteri
557	7	5.9	384	2	Q5FSL6_GLOUX	Q5fsl6 gluconobact	630	7	5.9	452	2	Q8RGR1_FUSNN	Q8rgr1 fusbacteri
558	7	5.9	388	1	Y2056_EDEBA	Q6mlf5 dbellovibri	631	7	5.9	453	2	Q9KP45_VIBCH	Q9kp45 vibrio chol
559	7	5.9	389	2	Q04178_BRACM	Q04178 brassica ca	632	7	5.9	454	2	Q4NXT4_9DELT	Q4nxt4 anaeromyxob
560	7	5.9	389	2	Q82567_ARATH	Q82567 arabidopsis	633	7	5.9	457	2	Q82M78_STRAW	Q82m78 streptomyc
561	7	5.9	389	2	Q9SP57_BRACM	Q9sp57 pseudomonas	634	7	5.9	459	2	Q7NVU0_CHNPK	Q7nvu0 chromobacte
562	7	5.9	389	2	Q912P7_PSEAE	Q912p7 pseudomonas	635	7	5.9	459	2	Q7U4V7_SYNPK	Q7u4v7 synchococc
563	7	5.9	390	2	Q9YDB2_AERPE	Q9ydb2 aeropyrum p	636	7	5.9	463	2	Q7Y3J1_MYCPA	Q7y3j1 mycobacteri
564	7	5.9	392	2	Q51RJ1_MAGGR	Q51rj1 magnaporthe	637	7	5.9	471	2	Q9PGM2_XYLPA	Q9pgm2 xylalla fas
565	7	5.9	392	2	Q8W1E9_BRARP	Q8wie9 brassica ra	638	7	5.9	471	2	Q7V8W3_PROMM	Q7v8w3 prochloroco
566	7	5.9	392	2	Q5YVF7_NOCFA	Q5yvf7 nocardia fa	639	7	5.9	473	2	Q8XHB7_NOCFA	Q8xhb7 nocardia ae
567	7	5.9	394	2	Q5AC31_CANAL	Q5ac31 candida alb	640	7	5.9	479	2	Q5YV45_NOCFA	Q5yv45 nocardia fa
568	7	5.9	394	2	Q5ACF2_CANAL	Q5acf2 candida alb	641	7	5.9	480	2	Q7U3Q4_SYNPK	Q7u3q4 synchococc
569	7	5.9	394	2	Q5H429_XANOR	Q5h429 xanthomonas	642	7	5.9	480	2	Q62165_BURMA	Q62165 burkholderi
570	7	5.9	396	2	Q579F3_DEIRA	Q579f3 bruceella ab	643	7	5.9	480	2	Q63RF9_BURPS	Q63rf9 burkholderi
571	7	5.9	396	2	Q9RS23_DEIRA	Q9rs23 deinococcus	644	7	5.9	483	2	Q75HF3_ORYSA	Q75hf3 oryza sativ
572	7	5.9	396	2	Q89QB8_BRAJA	Q89qb8 bradyrhizob	645	7	5.9	484	1	MURE_RHILO	Q98ka8 rhizobium l
573	7	5.9	396	2	Q937U1_BRUSU	Q937u1 bruceella su	646	7	5.9	486	2	Q4NTY0_9DELT	Q4nty0 anaeromyxob
574	7	5.9	396	2	Q8YD15_BRUME	Q8ydl5 bruceella me	647	7	5.9	488	2	Q7YVN3_9TRYP	Q7yvn3 trypanosoma
575	7	5.9	396	2	Q5YU77_NOCFA	Q5yuy7 nocardia fa	648	7	5.9	488	2	Q6DHU0_BRARE	Q6dhu0 brachydanio
576	7	5.9	398	2	Q7WHT1_BORBR	Q7wht1 bordetella	649	7	5.9	488	2	Q9OZ28_BRARE	Q9oz28 brachydanio
577	7	5.9	399	2	Q8BXV8_MOUSE	Q8bxv8 mus musculu	650	7	5.9	489	2	Q5B457_EXENI	Q5b457 aspergillus
578	7	5.9	401	2	Q6AU99_ORYSA	Q6au99 oryza sativ	651	7	5.9	492	2	Q5LUT5_SILPO	Q5lut5 silicibacte
579	7	5.9	403	2	Q8KYU7_9PROT	Q8kyu7 uncultured	652	7	5.9	493	1	PD12_CAREL	Q17770 caenorhabdi
580	7	5.9	405	2	Q69757_PSEFL	Q69757 pseudomonas	653	7	5.9	493	2	Q4U5F6_DICVI	Q4uf56 dictyocaulu
581	7	5.9	408	2	Q8XH17_CLOPE	Q8xh17 clostridium	654	7	5.9	493	2	Q618U9_CABVR	Q618u9 caenorhabdi
582	7	5.9	410	2	Q9HSM8_HALSA	Q9hsm8 halobacteri	655	7	5.9	493	2	Q6PST0_ANCCA	Q6pst0 ancylostoma
583	7	5.9	413	2	Q65CE8_9ACTO	Q65ce8 streptomyc	656	7	5.9	493	2	Q95PP0_OSTOS	Q95pp0 ostertagia
584	7	5.9	413	2	Q88HK7_PSEPK	Q88hk7 pseudomonas	657	7	5.9	494	2	Q524Y7_MAGGR	Q524y7 magnaporthe
585	7	5.9	413	2	Q7NZQ6_CHRVO	Q7nzq6 chromobacte	658	7	5.9	499	2	Q5YYI3_NOCFA	Q5yyi3 nocardia fa
586	7	5.9	414	2	Q9RMV4_VIBCH	Q9rmv4 vibrio chol	659	7	5.9	499	2	Q6N5K9_RHOPA	Q6n5k9 rhodopseudo
587	7	5.9	415	2	Q8DA59_VIBVY	Q8da59 vibrio vuln	660	7	5.9	499	2	Q74F16_GEOSL	Q74f16 geobacter s
588	7	5.9	415	2	Q7WK22_VIBVY	Q7wk22 vibrio vuln	661	7	5.9	500	2	Q8KY30_STRUC	Q8ky30 streptomyc
589	7	5.9	416	2	Q82136_STRAW	Q82136 streptomyc	662	7	5.9	500	2	Q8YPN6_XANSP	Q8ypn6 anabaena sp
590	7	5.9	417	2	Q34184_CEPNE	Q34184 cepaea nemo	663	7	5.9	504	2	Q7N1W5_GLOVI	Q7n1w5 gloeobacte
591	7	5.9	419	1	MURA_ACIBG	P33986 acinetobact	664	7	5.9	505	2	Q7N1W5_GLOVI	Q7n1w5 gloeobacte
592	7	5.9	419	1	Y1258_MYCTU	P64783 mycobacteri	665	7	5.9	510	2	Q9Y9C5_AERPE	Q9y9c5 aeropyrum p
593	7	5.9	419	1	Y1288_MYCBO	P64784 mycobacteri	666	7	5.9	510	2	Q9A748_CAUCR	Q9a748 caulobacter
594	7	5.9	419	2	Q5LX88_SILPO	Q5lx88 silicibacte	667	7	5.9	514	2	Q7N8B5_PHOLL	Q7n8b5 photorhabdu
595	7	5.9	421	2	Q4NJI2_9M1CC	Q4nji2 arthrobacte	668	7	5.9	515	2	Q9CA89_ARATH	Q9ca89 arbidopsis
596	7	5.9	421	2	Q4HA72_9DRIO	Q4ha72 deinococcus	669	7	5.9	519	2	Q4UTV7_CORYK	Q4jtv7 corynebacte
597	7	5.9	421	2	Q4UK37_RICPE	Q4uk37 rickettsia	670	7	5.9	521	2	Q6F2A2_MESFL	Q6f2a2 mesoplasma
598	7	5.9	421	2	Q7MAC1_WOLUS	Q7mac1 wolinnella s	671	7	5.9	525	2	Q5YW97_NOCFA	Q5yw97 nocardia fa
599	7	5.9	423	2	Q63MQ6_BURPS	Q63mq6 burkholderi	672	7	5.9	526	2	Q8PFU5_XANAC	Q8pfu5 xanthomonas
600	7	5.9	424	2	Q4NFE9_9M1CC	Q4nfe9 arthrobacte	673	7	5.9	529	2	Q4H578_9DEIO	Q4h578 deinococcus
601	7	5.9	426	2	Q9RLJ0_PSEAE	Q9rlj0 pseudomonas	674	7	5.9	530	2	Q9KP41_VIBCH	Q9kp41 vibrio chol
602	7	5.9	427	2	P95363_NEIGO	P95363 neisseria g	675	7	5.9	533	2	Q65QV0_MANNM	Q65qv0 pseudomonas
603	7	5.9	427	2	Q5F6G0_NEIG1	Q5f6g0 neisseria g	676	7	5.9	538	1	TNSE_ECOLI	Q7nq12 chromobacte
604	7	5.9	427	2	Q9JUS7_NEIMA	Q9juz7 neisseria m	677	7	5.9	538	2	Q7NQI2_CHRVO	Q7nq12 chromobacte
605	7	5.9	428	1	DAMX_ECO57	Q8x826 escherichia	678	7	5.9	543	2	Q5OTU2_IDILO	Q5ot2 idiomarina
606	7	5.9	428	1	DAMX_ECOLI	P11557 escherichia	679	7	5.9	543	2	Q92NE8_RHIME	Q92ne8 rhizobium m
607	7	5.9	428	2	Q8FCV8_ECOL6	Q8fcv8 escherichia	680	7	5.9	546	2	Q70FG1_PSEPU	Q70fg1 pseudomonas
608	7	5.9	429	2	Q4NG18_9M1CC	Q4ng18 arthrobacte	681	7	5.9	547	2	Q7NQI2_CHRVO	Q7nq12 chromobacte
609	7	5.9	429	2	Q7VVD4_BORPE	Q7vvd4 bordetella	682	7	5.9	551	2	Q981W0_RHILO	Q981w0 rhizobium l
610	7	5.9	429	2	Q7WKS4_BORBR	Q7wks4 bordetella	683	7	5.9	551	2	Q98CJ6_RHILO	Q98cj6 rhizobium l
611	7	5.9	430	2	Q4NXN0_9DELT	Q4nxn0 anaeromyxob	684	7	5.9	552	2	Q4UVQ9_XANCP	Q4uvq9 xanthomonas
612	7	5.9	430	2	Q83JA3_SHIFL	Q83ja3 shigella fl	685	7	5.9	552	2	Q8P8C8_XANCP	Q8p8c8 xanthomonas
613	7	5.9	431	1	SYH_NEIMA	Q9juz9 neisseria m	686	7	5.9	558	2	Q6LIA9_PHOPR	Q6lia9 photobacteri
614	7	5.9	431	1	SYH_NEIMA	Q9jux9 neisseria m	687	7	5.9	559	2	Q9HRT5_HALSA	Q9hrt5 halobacteri
615	7	5.9	431	2	Q5F9G8_NEIG1	Q5f9g8 neisseria g	688	7	5.9	563	2	Q8L800_FLABI	Q8l800 flavobact bi

689 7 5.9 563 2 Q5UPJ3 MIMIV Q5upj3 mimivirus.
690 7 5.9 564 2 Q7Q9Y0 ANOGA Q7q9y0 anopheles g
691 7 5.9 564 2 Q8L7Z9 SPIOL Q8l7z9 spinacia ol
692 7 5.9 568 2 Q7VZ95 BORPE Q7vz95 bordetella
693 7 5.9 568 2 Q7WAA6 BORPE Q7waa6 bordetella
694 7 5.9 568 2 Q7WJES BORBR Q7wjes bordetella
695 7 5.9 572 2 Q6L5F6 ORY8A Q6l5f6 oryza sativ
696 7 5.9 572 1 COK1B STRCO Q6l5f6 oryza sativ
697 7 5.9 582 2 Q8X8A7 ECO57 Q8x8a7 escherichia
698 7 5.9 583 2 Q8L7Z8 TOBAC Q8l7z8 nicotiana t
699 7 5.9 585 2 Q9KAF1 BACHD Q9kaf1 bacillus ha
700 7 5.9 588 2 Q5ASQ5 EMENI Q5asq5 aspergillus
701 7 5.9 589 2 Q9XBP3 MYXXA Q9xbp3 myxococcus
702 7 5.9 592 2 Q9VRN1 DROME Q9vrn1 drosophila
703 7 5.9 592 2 Q7YIH2 CLOAB Q7yih2 clostridium
704 7 5.9 594 2 Q7SUN9 HALDI Q7sun9 halictis di
705 7 5.9 594 2 Q86M37 HALDI Q86m37 halictis di
706 7 5.9 603 2 Q9RRP4 DEIRA Q9rrp4 deinococcus
707 7 5.9 605 2 P72607 SYN3 P72607 synchocyst
708 7 5.9 609 1 Y4PA RHIN P55610 rhizobium s
709 7 5.9 614 2 Q4IZJ5 AZOVI Q4izj5 azotobacter
710 7 5.9 617 2 Q59120 PYRHO Q59120 pyrococcus
711 7 5.9 626 2 Q9NXC4 HUMAN Q9nxc4 homo sapien
712 7 5.9 627 2 Q4C4V0 LEIMA Q4c4v0 leishmania
713 7 5.9 627 2 Q5WBV2 BACSK Q5wbv2 bacillus cl
714 7 5.9 627 2 Q8VBV5 MOUSE Q8vbv5 mus musculu
715 7 5.9 629 2 Q6LTG7 PHOPR Q6ltg7 photobacter
716 7 5.9 632 2 Q5B5C3 EMENI Q5b5c3 aspergillus
717 7 5.9 645 2 Q9RU98 DEIRA Q9ru98 deinococcus
718 7 5.9 647 2 Q8H6T8 HUMAN Q8h6t8 homo sapien
719 7 5.9 647 2 Q6P5W5 HUMAN Q6p5w5 homo sapien
720 7 5.9 657 2 Q5Z0R6 NOCPA Q5z0r6 nocardia fa
721 7 5.9 665 2 Q8ZBL8 YERPE Q8zbl8 yersinia pe
722 7 5.9 665 2 Q66EF2 YERPS Q66ef2 yersinia ps
723 7 5.9 667 2 Q8ZUL8 PYRAE Q8zul8 pyrobaculum
724 7 5.9 669 2 Q584R6 9TRYP Q584r6 trypanosoma
725 7 5.9 675 1 H2PA XANAC Q8ph20 xanthomonas
726 7 5.9 686 2 Q8A4X3 BACTN Q8a4x3 bacteroides
727 7 5.9 689 2 Q8D1A4 YERPE Q8d1a4 yersinia pe
728 7 5.9 691 2 Q9AF95 9RHO Q9af95 pseudomonas
729 7 5.9 694 2 Q4LLY8 9BURK Q4lly8 burkholderi
730 7 5.9 704 2 Q6FTC9 CANGA Q6ftc9 candida gla
731 7 5.9 707 2 Q5WV32 LEGPL Q5wv32 legionella
732 7 5.9 707 2 Q5XK34 LEGPA Q5xk34 legionella
733 7 5.9 712 2 Q9K3K0 STRCO Q9k3k0 streptomyce
734 7 5.9 714 2 Q4HHU9 CAMCO Q4hhu9 campylobact
735 7 5.9 714 2 Q72FC0 DESVH Q72fc0 desulfovibr
736 7 5.9 715 2 Q5ZT29 LSGPH Q5zt29 legionella
737 7 5.9 728 2 QANT17 9BELT Qant17 anaeromyxob
738 7 5.9 729 2 Q8GRP8 ORISA Q8grp8 oryza sativ
739 7 5.9 729 2 Q8VS76 BACTN Q8vs76 bacteroides
740 7 5.9 729 2 Q5HVE9 CAMJR Q5hve9 campylobact
741 7 5.9 729 2 Q9PHP5 CAMJBE Q9php5 campylobact
742 7 5.9 733 2 Q92XH9 RHIME Q92xh9 rhizobium m
743 7 5.9 735 2 Q4UWV7 XANCP Q4uwv7 xanthomonas
744 7 5.9 735 2 Q8P786 XANCP Q8p786 xanthomonas
745 7 5.9 737 2 Q4UQ53 XANCP Q4uq53 xanthomonas
746 7 5.9 737 2 Q8P4J8 XANCP Q8p4j8 xanthomonas
747 7 5.9 740 2 Q83118 TREPA Q83118 treponema p
748 7 5.9 755 2 Q51AK4 ENTHI Q51ak4 entamoeba h
749 7 5.9 758 2 Q7NDC5 GLOVI Q7ndc5 gloebacter
750 7 5.9 766 2 Q8C4W5 MOUSE Q8c4w5 mus musculu
751 7 5.9 769 2 Q8PQ37 XANAC Q8pq37 xanthomonas
752 7 5.9 774 2 Q8AAK5 BACTN Q8aak5 bacteroides
753 7 5.9 774 2 Q7WAE7 BORPA Q7wae7 bordetella
754 7 5.9 778 2 Q7WJJO BORBR Q7wjjo bordetella
755 7 5.9 778 2 Q7VVE8 BORPE Q7vve8 bordetella
756 7 5.9 785 2 Q8NPP6 CORGL Q8npp6 corynebacte
757 7 5.9 790 2 Q742J2 MYCPA Q742j2 mycobacteri
758 7 5.9 797 1 CTPE MYCBO P0a505 mycobacteri
759 7 5.9 797 1 CTPE MYCBO P0a505 mycobacteri
760 7 5.9 815 1 SL9A1 HUMAN P19634 homo sapien
761 7 5.9 815 2 Q5VU07 HUMAN Q5vu07 homo sapien

762 7 5.9 816 1 SL9A1 RABIT 762
763 7 5.9 817 1 SL9A1 BOVIN 763
764 7 5.9 818 1 SL9A1 PIG 764
765 7 5.9 820 1 SL9A1 MOUSE 765
766 7 5.9 820 1 SL9A1 RAT 766
767 7 5.9 820 1 Q80X31 MOUSE 767
768 7 5.9 823 2 Q7TSV2 MOUSE 768
769 7 5.9 844 2 Q89ZN9 BACTN 769
770 7 5.9 870 2 Q8X316 ECO57 770
771 7 5.9 875 2 Q8T7V7 TRYCR 771
772 7 5.9 881 2 Q687F3 9VIRU 772
773 7 5.9 881 2 Q9EVE1 ECO57 773
774 7 5.9 881 2 Q91537 PSEAE 774
775 7 5.9 884 2 Q8X327 ECO57 775
776 7 5.9 900 2 Q7QET0 ANOGA 776
777 7 5.9 914 2 Q53964 STRCO 777
778 7 5.9 917 2 Q8T7V6 TRYCR 778
779 7 5.9 925 2 Q15637 TRYCR 779
780 7 5.9 937 2 Q8TSM9 HUMAN 780
781 7 5.9 961 2 Q7VN93 HAEDU 781
782 7 5.9 974 1 ATXA LEIDO 782
783 7 5.9 974 1 ATXB LEIDO 783
784 7 5.9 974 2 Q4QDN8 LEIMA 784
785 7 5.9 974 2 Q4QDN7 LEIMA 785
786 7 5.9 975 2 Q8MZZ9 NEUCR 786
787 7 5.9 984 2 Q5TBP2 HUMAN 787
788 7 5.9 985 2 Q7SHP3 NEUCR 788
789 7 5.9 986 2 Q4HBU4 9DEIO 789
790 7 5.9 1005 1 EVC MOUSE 790
791 7 5.9 1009 2 Q7M912 WOLSU 791
792 7 5.9 1053 2 Q4R5S1 MACPA 792
793 7 5.9 1099 2 Q7TUJ6 PROMM 793
794 7 5.9 1200 2 Q54HL3 DICDI 794
795 7 5.9 1219 2 Q53785 STRAU 795
796 7 5.9 1225 2 Q82DD9 STRAW 796
797 7 5.9 1255 2 Q6M414 CORGL 797
798 7 5.9 1268 2 Q9KZL1 STRCO 798
799 7 5.9 1302 2 Q8C8X7 MOUSE 799
800 7 5.9 1338 2 Q4P748 USTWA 800
801 7 5.9 1343 2 Q06635 9ALPH 801
802 7 5.9 1343 2 Q4LDG6 9ALPH 802
803 7 5.9 1354 2 Q5E034 VIBF1 803
804 7 5.9 1385 2 Q65565 9ALPH 804
805 7 5.9 1385 2 Q77CC1 9ALPH 805
806 7 5.9 1391 2 Q6X230 9ALPH 806
807 7 5.9 1392 2 Q9XER9 ARATH 807
808 7 5.9 1408 2 Q6X125 9ALPH 808
809 7 5.9 1446 1 IE18 PRVKA 809
810 7 5.9 1446 2 Q5PP75 9ALPH 810
811 7 5.9 1454 2 Q8JLE3 9ALPH 811
812 7 5.9 1461 1 IE18 PRVIF 812
813 7 5.9 1634 2 Q51IK0 MAGGR 813
814 7 5.9 1764 2 Q5BCS0 EMENI 814
815 7 5.9 1820 2 Q6C604 YARLI 815
816 7 5.9 2056 2 Q8CHF3 MOUSE 816
817 7 5.9 2058 2 Q5T4S9 HUMAN 817
818 7 5.9 2276 2 Q75050 HUMAN 818
819 7 5.9 2954 2 Q4FWP4 LEIMA 819
820 7 5.9 3032 1 POLG HCJV6 820
821 7 5.9 3032 2 Q99IE4 9HEPC 821
822 7 5.9 3032 2 Q9Q9A9 9HEPC 822
823 7 5.9 3033 2 Q99IB3 9HEPC 823
824 7 5.9 3033 2 Q99IB6 9HEPC 824
825 7 5.9 3033 2 Q99IB7 9HEPC 825
826 7 5.9 3033 2 Q99IB8 9HEPC 826
827 7 5.9 3033 2 Q91Z2 9HEPC 827
828 7 5.9 3033 2 Q91Z4 9HEPC 828
829 7 5.9 3033 2 Q91Z5 9HEPC 829
830 7 5.9 3033 2 Q91Z6 9HEPC 830
831 7 5.9 3033 2 Q9Q9A7 9HEPC 831
832 7 5.9 3033 2 Q9Q9A8 9HEPC 832
833 7 5.9 3033 2 Q9Q9B0 9HEPC 833
834 7 5.9 3033 2 Q9QF35 9HEPC 834

P23791 oryctolegus
Q28036 bos taurus
P48762 sus scrofa
Q61165 mus musculus
P26431 rattus norv
Q80x31 mus musculus
Q7cav2 mus musculus
Q89zn9 bacteroides
Q8x316 escherichia
Q8t7v7 trypanosoma
Q687f3 bacterioph
Q8eyel escherichia
Q91537 pseudomonas
Q8x327 escherichia
Q7qeto anopheles g
Q53964 streptomyce
Q8t7v6 trypanosoma
Q15637 trypanosoma
Q8tbn9 homo sapien
Q7vn93 haemophilus
P11718 leishmania
P12522 leishmania
Q4qdn8 leishmania
Q4qdn7 leishmania
Q8wz9 neurospora
Q5tbp2 homo sapien
Q4shp3 neurospora
Q4hbua4 deinococcus
P57680 mus musculus
Q7m912 wolinnella s
Q4r581 macaca fasc
Q7tj6 prochloroco
Q54hl3 dictyosteli
Q53785 streptomyce
Q82dd9 streptomyce
Q6m414 corynebacte
Q9kz11 streptomyce
Q8c8x7 mus musculus
Q4p748 utillago ma
Q06635 bovine herp
Q4ldg6 bovine herp
Q5e034 vibrio fasc
Q85565 bovine herp
Q77cc1 bovine herp
Q6x230 bovine herp
Q9xer9 arabidopsais
Q6x125 bovine herp
P33479 pseudorabie
Q5pp75 suid herpes
Q8jle3 suid herpes
P11675 pseudorabie
Q5lik0 magnaporthe
Q5bc90 aspergillus
Q6c604 yarrowia li
Q8chf3 mus musculus
Q5c489 homo sapien
Q75050 homo sapien
Q4fwp4 leishmania
P26660 h genome po
Q99ib4 hepatitis c
Q99ib3 hepatitis c
Q99ib6 hepatitis c
Q99ib7 hepatitis c
Q99ib8 hepatitis c
Q91za2 hepatitis c
Q91za4 hepatitis c
Q91za5 hepatitis c
Q91za6 hepatitis c
Q9Q9a7 hepatitis c
Q9Q9a8 hepatitis c
Q9Q9b0 hepatitis c
Q9Qf35 hepatitis c

835	7	5.9	3033	2	Q9QAX1_9HEPC	Q9QAX1	hepatitis c	908	6	5.1	60	2	Q9MC16_9CAUD	Q9mc16	bacterioph
836	7	5.9	3033	2	Q7T7H9_9HEPC	Q7t7h9	hepatitis c	909	6	5.1	60	2	Q88IN6_PSRPK	Q88in6	pseudomonas
837	7	5.9	3033	2	Q7T7I0_9HEPC	Q7t7i0	hepatitis c	910	6	5.1	61	1	AKH2_L0CMI	P08379	locusta mig
838	7	5.9	3033	2	Q7T7I5_9HEPC	Q7t7i5	hepatitis c	911	6	5.1	61	2	Q4TJ62_TETNG	Q4tj62	tetradon n
839	7	5.9	3033	2	Q7T7I6_9HEPC	Q7t7i6	hepatitis c	912	6	5.1	62	2	Q7UK05_RHOBA	Q7uk05	rhodopirell
840	7	5.9	3033	2	Q7T7I9_9HEPC	Q7t7i9	hepatitis c	913	6	5.1	62	2	Q63372_9ACTI	O63372	acipenser o
841	7	5.9	3033	2	Q7T7J0_9HEPC	Q7t7j0	hepatitis c	914	6	5.1	62	2	Q63373_9ACTI	O63373	acipenser o
842	7	5.9	3033	2	Q7T7J1_9HEPC	Q7t7j1	hepatitis c	915	6	5.1	62	2	Q63375_ACISA	O63375	acipenser s
843	7	5.9	3033	2	Q7T7J2_9HEPC	Q7t7j2	hepatitis c	916	6	5.1	62	2	Q63386_ACIDA	O63386	acipenser d
844	7	5.9	3033	2	Q91Z23_9HEPC	Q91z23	hepatitis c	917	6	5.1	62	2	Q63374_ACIPF	O63374	acipenser p
845	7	5.9	3034	2	Q5XNK5_9HEPC	Q5xnk5	hepatitis c	918	6	5.1	62	2	Q63370_ACINA	O63370	acipenser n
846	7	5.9	3037	2	Q68749_9HEPC	Q68749	hepatitis c	919	6	5.1	62	2	Q63371_ACINU	O63371	acipenser n
847	7	5.9	5183	2	Q8TDN5_HUMAN	Q8tdn5	homo sapien	920	6	5.1	63	2	Q6EQV3_ORYSA	O6eqv3	oryza sativ
848	7	5.9	5183	2	Q5TAS7_HUMAN	Q5tas7	homo sapien	921	6	5.1	63	2	Q4HDV7_CAMUP	Q4hdv7	campylobact
849	7	5.9	5644	2	Q93NX8_9ACTO	Q93nx8	streptomyce	922	6	5.1	63	2	Q4HR59_CAMUP	Q4hr59	campylobact
850	7	5.9	7349	2	Q8GMP2_STRAZ	Q8gmp2	streptomyce	923	6	5.1	63	2	Q7VTJ1_BORPA	Q7vtj1	bordetella
851	18	5.1	18	2	Q7M2Q1_MACFA	Q7m2q1	macaca fasc	924	6	5.1	63	2	Q7W7H0_BORPA	Q7w7h0	bordetella
852	6	5.1	19	2	Q84274_HPV25	Q84274	human papil	925	6	5.1	63	2	Q7WKV9_BORBR	Q7wkv9	bordetella
853	6	5.1	20	2	Q9QXK7_9MURI	Q9qxk7	rattus sp.	926	6	5.1	63	2	Q8SDJ6_APOSY	Q8sdj6	apodemus sy
854	6	5.1	22	2	Q9QXK6_9MURI	Q9qxk6	rattus sp.	927	6	5.1	63	2	Q52262_9VIRU	Q52262	alternanthe
855	6	5.1	29	2	Q4RAB0_TETNG	Q4rab0	tetradon n	928	6	5.1	63	2	Q52265_9VIRU	Q52265	alternanthe
856	6	5.1	30	2	Q9QV18_9MURI	Q9qv18	rattus sp.	929	6	5.1	63	2	Q52267_9VIRU	Q52267	alternanthe
857	6	5.1	30	2	Q9QV19_9MURI	Q9qv19	rattus sp.	930	6	5.1	64	2	Q4TSP8_9SPHN	Q4tsf8	erythrobact
858	6	5.1	33	2	Q4UVH7_XANCP	Q4uvh7	xanthomonas	931	6	5.1	65	2	Q4JB65_SULAC	Q4jb65	sulfolobus
859	6	5.1	33	2	Q8PBK6_XANCP	Q8pbk6	xanthomonas	932	6	5.1	65	2	Q92N90_RHIME	Q92n90	rhizobium m
860	6	5.1	36	2	Q7RSC3_PLAYO	Q7rsc3	plasmodium	933	6	5.1	65	2	Q4S2D5_TETNG	Q4s2d5	tetradon n
861	6	5.1	36	2	Q8PK38_XANAC	Q8pk38	xanthomonas	934	6	5.1	67	2	Q0S613_PSRSP	O0s613	pseudomonas
862	6	5.1	41	2	Q81247_9HEPC	Q81247	hepatitis c	935	6	5.1	67	2	Q7ME50_VIBVY	Q7me50	vibrio vuln
863	6	5.1	41	2	Q81248_9HEPC	Q81248	hepatitis c	936	6	5.1	68	2	Q4HCD9_9DSIO	Q4hcd9	deinococcus
864	6	5.1	41	2	Q81249_9HEPC	Q81249	hepatitis c	937	6	5.1	68	2	Q9QXK6_9POKV	Q9qxk6	rabbit fibr
865	6	5.1	41	2	Q81251_9HEPC	Q81251	hepatitis c	938	6	5.1	69	1	YJBJ_ECOL6	Yjbj	ecoli
866	6	5.1	41	2	Q81252_9HEPC	Q81252	hepatitis c	939	6	5.1	69	1	YJBJ_ECOL6	Yjbj	ecoli
867	6	5.1	42	2	Q63Q33_BURPS	Q63q33	burkholderi	940	6	5.1	69	1	YJBJ_ECOL6	Yjbj	ecoli
868	6	5.1	42	2	Q63Q33_BURPS	Q63q33	burkholderi	941	6	5.1	69	1	YJBJ_ECOL6	Yjbj	ecoli
869	6	5.1	42	2	Q62DD7_BURMA	Q62dd7	burkholderi	942	6	5.1	69	2	Q9CP44_PASMU	Q9cp44	pasteurella
870	6	5.1	45	2	Q5LPR6_SILPO	Q5lpr6	silicibacte	943	6	5.1	69	2	Q9S570_ASTDO	Q9s570	asthenes do
871	6	5.1	45	2	Q8SKD4_ACISU	Q8skd4	acipenser s	944	6	5.1	69	2	Q9S560_9FURN	Q9s560	cranioloeuca
872	6	5.1	47	2	Q4RCR3_TETNG	Q4rcr3	tetradon n	945	6	5.1	69	2	Q9S563_9FURN	Q9s563	cranioloeuca
873	6	5.1	50	2	Q94VL8_ACIGU	Q94vl8	acipenser g	946	6	5.1	69	2	Q9S559_9FURN	Q9s559	cranioloeuca
874	6	5.1	50	2	Q63380_ACIME	Q63380	acipenser m	947	6	5.1	69	2	Q9S568_9FURN	Q9s568	cranioloeuca
875	6	5.1	50	2	Q63381_ACIBE	Q63381	acipenser b	948	6	5.1	69	2	Q9S558_9FURN	Q9s558	cranioloeuca
876	6	5.1	50	2	Q63392_POISP	Q63392	polyodon sp	949	6	5.1	69	2	Q9S564_9FURN	Q9s564	cranioloeuca
877	6	5.1	50	2	Q63376_9ACTI	Q63376	psephurus g	950	6	5.1	69	2	Q9S562_9FURN	Q9s562	cranioloeuca
878	6	5.1	50	2	Q63385_ACISU	Q63385	acipenser s	951	6	5.1	69	2	Q9S571_9FURN	Q9s571	cranioloeuca
879	6	5.1	50	2	Q94VM0_ACINA	Q94vm0	acipenser n	952	6	5.1	69	2	Q9S584_9FURN	Q9s584	cranioloeuca
880	6	5.1	50	2	Q94VL7_9ACTI	Q94vl7	pseudoscaph	953	6	5.1	69	2	Q9S565_9FURN	Q9s565	cranioloeuca
881	6	5.1	50	2	Q63377_ACIRT	Q63377	acipenser r	954	6	5.1	69	2	Q9S566_9FURN	Q9s566	cranioloeuca
882	6	5.1	50	2	Q63387_ACIFU	Q63387	acipenser f	955	6	5.1	69	2	Q9S567_9FURN	Q9s567	cranioloeuca
883	6	5.1	50	2	Q63382_ACIBR	Q63382	acipenser b	956	6	5.1	69	2	Q9S561_9FURN	Q9s561	cranioloeuca
884	6	5.1	50	2	Q94VM4_ACIPF	Q94vm4	acipenser p	957	6	5.1	70	2	Q74K20_LACTO	Q74k20	lactobacill
885	6	5.1	50	2	Q94VM2_ACISC	Q94vm2	acipenser s	958	6	5.1	70	2	Q9HND9_HALSA	Q9hnd9	halobacteri
886	6	5.1	50	2	Q94VL9_ACITR	Q94vl9	acipenser t	959	6	5.1	71	2	Q4Z987_9CAUD	Q4z987	staphylococ
887	6	5.1	50	2	Q94VL6_9ACTI	Q94vl6	scaphirhync	960	6	5.1	71	2	Q4LRG1_9BURK	Q4lrg1	burkholderi
888	6	5.1	50	2	Q94VM3_ACIBE	Q94vm3	acipenser b	961	6	5.1	71	2	Q6YQ85_ONYPE	Q6yq85	onion yello
889	6	5.1	50	2	Q63384_9ACTI	Q63384	huso dauric	962	6	5.1	71	2	Q6YQ85_ONYPE	Q6yq85	onion yello
890	6	5.1	50	2	Q63369_ACIMI	Q63369	acipenser m	963	6	5.1	73	2	Q8MVK3_9ASCI	Q8mvk3	boltenia vi
891	6	5.1	50	2	Q63388_ACIGU	Q63388	acipenser g	964	6	5.1	74	2	Q8MVK3_9ASCI	Q8mvk3	boltenia vi
892	6	5.1	50	2	Q94VM1_ACIMI	Q94vm1	acipenser m	965	6	5.1	74	2	Q8B6S0_9AVES	Q8b6s0	mullerornis
893	6	5.1	50	2	Q63378_ACISC	Q63378	acipenser s	966	6	5.1	75	2	Q6K1V7_ORYSA	Q6k1v7	oryza sativ
894	6	5.1	50	2	Q7J7T5_9ACTI	Q7j7t5	scaphirhync	967	6	5.1	75	2	Q4TLK4_9SPHN	Q4tlk4	erythrobact
895	6	5.1	50	2	Q63383_9ACTI	Q63383	huso huso (968	6	5.1	75	2	Q9S699_9AVES	Q9s699	neomorphus
896	6	5.1	50	2	Q63379_ACIST	Q63379	acipenser s	969	6	5.1	76	2	Q41497_GIBZEL	Q41497	gibberella
897	6	5.1	50	2	Q63393_9ACTI	Q63393	pseudoscaph	970	6	5.1	77	2	Q75E13_ASHGO	Q75e13	ashbya goss
898	6	5.1	51	1	Y14_BPT7	P03791	bacterioph	971	6	5.1	77	2	Q9RCD9_XANCA	Q9rcd9	xanthomonas
899	6	5.1	51	2	Q3E509_BPT7	Q3e509	bacterioph	972	6	5.1	77	2	Q9RGY6_LACAC	Q9rgy6	lactobacill
900	6	5.1	51	2	Q6WY53_BPT7	Q6wy53	bacterioph	973	6	5.1	77	2	Q92K74_RHIME	Q92k74	rhizobium m
901	6	5.1	51	2	Q6WYF7_BPT7	Q6wyf7	bacterioph	974	6	5.1	77	2	Q89YC1_BRAJA	Q89yc1	bradyrhizob
902	6	5.1	51	2	Q6WYR3_BPT7	Q6wyr3	bacterioph	975	6	5.1	77	2	Q98S7_RHILQ	Q98s7	rhizobium l
903	6	5.1	52	2	Q4TFP6_TETNG	Q4tfp6	tetradon n	976	6	5.1	78	2	Q7YD61_RABIT	Q7yd61	oryctolagus
904	6	5.1	54	2	Q4YM45_PLABE	Q4ym45	plasmodium	977	6	5.1	78	2	Q936T9_PSESP	Q936t9	pseudomonas
905	6	5.1	55	2	Q4WTN0_9DELT	Q4wtm0	anaeromyxob	978	6	5.1	79	2	Q56UF8_LYNST	Q56uf8	lynaea sta
906	6	5.1	58	2	Q5L3F0_GEOKA	Q5l3f0	geobacillus	979	6	5.1	80	1	GCH1_MUCHA	Gch1	mucuna hass
907	6	5.1	59	2	Q4T2U0_TETNG	Q4t2u0	tetradon n	980	6	5.1	80	2	Q5XNS5_ANOGA	Q5xns5	anopheles g

981	6	5.1	80	2	Q743K5_MYCPA	Q743k5_mycobacteri	1054	6	5.1	93	2	Q5BMW4_XANCA	Q5bmw4_xanthomonas
982	6	5.1	81	2	Q5ZB39_ORYSA	Q5zb39_oryza sativ	1055	6	5.1	93	2	Q5BMX2_XANCA	Q5bmX2_xanthomonas
983	6	5.1	81	2	Q6N83_ORYSA	Q6n83_oryza sativ	1056	6	5.1	93	2	Q8M240_FRANA	Q8m240_fraomyces nat
984	6	5.1	81	2	Q06779_MYCTU	Q06779_mycobacteri	1057	6	5.1	93	2	Q7YGG0_9NEOB	Q7ygg0_bufo woodho
985	6	5.1	81	2	Q7U117_MYCBO	Q7u117_mycobacteri	1058	6	5.1	94	2	Q97ZB1_ECULI	Q97ze1_escherichia
986	6	5.1	81	2	Q820B9_ANASP	Q820b9_anabaena sp	1059	6	5.1	94	2	Q9F576_ECULI	Q9f576_escherichia
987	6	5.1	81	2	Q98MZ8_RHILO	Q98mz8_rhizobium l	1060	6	5.1	94	2	Q8FLI3_COREF	Q8fli3_corynebacte
988	6	5.1	82	2	Q98694_GAVES	Q98694_cuculus can	1061	6	5.1	94	2	Q62CJ8_BURMA	Q62cj8_burkholderi
989	6	5.1	82	2	Q99697_GEOCA	Q99697_geococcyx c	1062	6	5.1	94	2	Q20499_9PASS	Q20499_andropadus
990	6	5.1	83	2	Q9HRM2_HALSA	Q9hrw2_halobacteri	1063	6	5.1	94	2	Q20497_9PASS	Q20497_andropadus
991	6	5.1	83	2	Q69KP4_ORYSA	Q69kp4_oryza sativ	1064	6	5.1	94	2	Q20501_9PASS	Q20501_andropadus
992	6	5.1	83	2	Q84QX9_ORYSA	Q84qx9_oryza sativ	1065	6	5.1	94	2	Q9T6P3_9TELE	Q9t6p3_acrossochei
993	6	5.1	83	2	Q84KGS_MYCTU	Q84kg8_mycobacteri	1066	6	5.1	94	2	Q20491_9PASS	Q20491_andropadus
994	6	5.1	83	2	Q7VVF5_BORPE	Q7vvf5_bordetella	1067	6	5.1	94	2	Q20505_9PASS	Q20505_andropadus
995	6	5.1	83	2	Q7WB93_BORPA	Q7wb93_bordetella	1068	6	5.1	94	2	Q20489_9PASS	Q20489_andropadus
996	6	5.1	83	2	Q7WMR1_BORBR	Q7wmr1_bordetella	1069	6	5.1	94	2	Q20509_9PASS	Q20509_andropadus
997	6	5.1	83	2	Q9A171_STRPY	Q9a171_streptomyces	1070	6	5.1	94	2	Q20503_9PASS	Q20503_andropadus
998	6	5.1	83	2	Q9S2U4_STRPY	Q9s2u4_streptomyces	1071	6	5.1	94	2	Q20495_9PASS	Q20495_andropadus
999	6	5.1	83	2	Q8P292_STRP8	Q8p292_streptococ	1072	6	5.1	94	2	Q20493_9PASS	Q20493_andropadus
1000	6	5.1	83	2	Q608F1_METCA	Q608f1_methylococcyx	1073	6	5.1	94	2	Q20511_9PASS	Q20511_phyllastrep
1001	6	5.1	83	2	Q99696_GAVES	Q99696_dromococcyx	1074	6	5.1	94	2	Q21852_9PASS	Q21852_andropadus
1002	6	5.1	83	2	Q99700_GAVES	Q99700_phaenocopha	1075	6	5.1	94	2	Q21782_9PASS	Q21782_andropadus
1003	6	5.1	83	2	Q99703_GAVES	Q99703_tapera naev	1076	6	5.1	95	2	Q97516_9PRIM	Q97516_gorilla gor
1004	6	5.1	83	2	Q99690_GAVES	Q99690_cercococcyx	1077	6	5.1	95	2	Q9T6P1_9TELE	Q9t6p1_acrossochei
1005	6	5.1	83	2	Q99695_GAVES	Q99695_cuculus pol	1078	6	5.1	95	2	Q9T6N9_9TELE	Q9t6n9_acrossochei
1006	6	5.1	83	2	Q99701_PIACA	Q99701_playa cayan	1079	6	5.1	95	2	Q20507_9PASS	Q20507_andropadus
1007	6	5.1	83	2	Q99691_GAVES	Q99691_chrysococcy	1080	6	5.1	95	2	Q9T6P0_9TELE	Q9t6p0_acrossochei
1008	6	5.1	85	2	Q874Y3_PODAN	Q874y3_chrysospora a	1081	6	5.1	95	2	Q9T6P2_9TELE	Q9t6p2_acrossochei
1009	6	5.1	85	2	Q6WRM0_HAPAL	Q6wrw0_hafnia alve	1082	6	5.1	95	2	Q9G6E1_CORCN	Q9g6e1_corythaiXoi
1010	6	5.1	85	2	Q4T498_TRING	Q4t498_tetradodon n	1083	6	5.1	95	2	Q9T6P4_9TELE	Q9t6p4_acrossochei
1011	6	5.1	86	2	Q526T8_MAGGR	Q526t8_magnaporthe	1084	6	5.1	95	2	Q9PTA9_9TELE	Q9pta9_acrossochei
1012	6	5.1	86	2	Q9WM82_LACIC	Q9wm82_lactococcus	1085	6	5.1	95	2	Q9T4J9_9TELE	Q9t4j9_acrossochei
1013	6	5.1	86	2	Q9REX9_9LACT	Q9rex9_lactococcus	1086	6	5.1	95	2	Q9T4J1_9TELE	Q9t4j1_acrossochei
1014	6	5.1	86	2	Q34015_CHAAL	Q34015_charadrius	1087	6	5.1	95	2	Q9T3H1_9TELE	Q9t3h1_acrossochei
1015	6	5.1	86	2	Q30220_9CHAR	Q30220_charadrius	1088	6	5.1	95	2	Q4TVV5_COLLI	Q4tvv5_columba liv
1016	6	5.1	86	2	Q30214_9CHAR	Q30214_charadrius	1089	6	5.1	95	2	Q5WSP4_LEGPL	Q5wsp4_legionella
1017	6	5.1	86	2	Q30216_9CHAR	Q30216_charadrius	1090	6	5.1	96	2	Q5ZRG8_LEGPH	Q5zrg8_legionella
1018	6	5.1	86	2	Q34026_9CHAR	Q34026_charadrius	1091	6	5.1	96	2	Q6AGT8_LEIXX	Q6agt8_leifsonia x
1019	6	5.1	86	2	Q34013_9CHAR	Q34013_charadrius	1092	6	5.1	96	2	Q47867_OTITA	Q47867_otitis tarda
1020	6	5.1	86	2	Q30208_9CHAR	Q30208_charadrius	1093	6	5.1	96	2	Q4W8Y1_ASPFU	Q4w8y1_aspergillus
1021	6	5.1	86	2	Q30228_VANCH	Q30228_vanellus ch	1094	6	5.1	97	2	Q9XN52_GIRCA	Q9xn52_giraffa cam
1022	6	5.1	86	2	Q30222_9CHAR	Q30222_eleyornis	1095	6	5.1	97	2	Q94UG8_LUTCA	Q94ug8_lutra canad
1023	6	5.1	86	2	Q30218_9CHAR	Q30218_charadrius	1096	6	5.1	97	2	Q5N0A2_SYNP6	Q5n0a2_synechococ
1024	6	5.1	86	2	Q30210_CHAPA	Q30210_charadrius	1097	6	5.1	97	2	Q5BC27_EMENI	Q5bc27_espergillus
1025	6	5.1	86	2	Q34271_CHAVO	Q34271_charadrius	1098	6	5.1	98	2	Q912H5_PSEAE	Q912h5_pseudomonas
1026	6	5.1	86	2	Q30224_9CHAR	Q30224_oreopholus	1099	6	5.1	98	2	Q9YWS7_9HIV1	Q9yws7_human immun
1027	6	5.1	86	2	Q30212_9CHAR	Q30212_charadrius	1100	6	5.1	98	2	Q9YWT5_9HIV1	Q9ywt5_human immun
1028	6	5.1	86	2	Q30226_9CHAR	Q30226_thinornis r	1101	6	5.1	98	2	Q47389_CROEL	Q47389_crocidura e
1029	6	5.1	87	2	Q6ZJQ5_ORYSA	Q6zjq5_oryza sativ	1102	6	5.1	99	2	Q47387_CROML	Q47387_crocidura m
1030	6	5.1	88	2	Q4JD21_PANLE	Q4jd21_panthera le	1103	6	5.1	99	2	Q47391_CROLB	Q47391_crocidura l
1031	6	5.1	88	2	Q9GC00_BURPA	Q9gc00_burranyas pa	1104	6	5.1	99	2	Q47388_CRORH	Q47388_crocidura r
1032	6	5.1	88	2	Q39897_9HEPC	Q39897_hepatitis c	1105	6	5.1	99	2	Q47385_CROFU	Q47385_crocidura f
1033	6	5.1	88	2	Q39914_9HEPC	Q39914_hepatitis c	1106	6	5.1	99	2	Q7J6C2_CROFS	Q7j6c2_crocidura r
1034	6	5.1	89	2	Q48302_HALSA	Q48302_halobacteri	1107	6	5.1	99	2	Q5B0U1_SOROR	Q5b0u1_sorex ornat
1035	6	5.1	89	2	Q39895_9HEPC	Q39895_hepatitis c	1108	6	5.1	99	2	Q56Z06_ARATH	Q56z06_arabidopsis
1036	6	5.1	90	2	Q39911_9HEPC	Q39911_hepatitis c	1109	6	5.1	99	2	Q9XKY0_EUPSE	Q9xky0_eupodotis s
1037	6	5.1	91	2	Q97517_9PRIM	Q97517_gorilla gor	1110	6	5.1	99	2	Q9XKY1_9GRUI	Q9xky1_ardeotis ko
1038	6	5.1	91	2	P96270_MYCTU	P96270_mycobacteri	1111	6	5.1	99	2	Q9XKY3_ANTVI	Q9xky3_anthropoide
1039	6	5.1	91	2	Q7U211_MYCBO	Q7u211_mycobacteri	1112	6	5.1	99	2	Q9XKY2_9GRUI	Q9xky2_fulica cris
1040	6	5.1	91	2	Q8M242_MASCO	Q8m242_mastomys co	1113	6	5.1	100	1	YOG4_CAREL	Y34613_caenorhabdi
1041	6	5.1	91	2	Q8M239_PRANA	Q8m239_praomyces nat	1114	6	5.1	100	2	Q5XKU2_9CUCU	Q5xku2_apriona gpr
1042	6	5.1	92	2	Q8XY9_DROME	Q8xy9_drosophila	1115	6	5.1	100	2	Q7NYH6_CHRVO	Q7nyh6_chromobacte
1043	6	5.1	92	2	Q54078_SACER	Q54078_saccharopol	1116	6	5.1	100	2	Q72945_9HIV1	Q72945_human immun
1044	6	5.1	92	2	Q8VKL4_MYCTU	Q8vkl4_mycobacteri	1117	6	5.1	101	2	Q9B864_SOROR	Q9b864_sorex ornat
1045	6	5.1	92	2	Q5PBC2_ANAMM	Q5pbc2_anaplasma m	1118	6	5.1	101	2	Q9B865_SORVA	Q9b865_sorex vagra
1046	6	5.1	92	2	Q73T07_MYCPA	Q73t07_mycobacteri	1119	6	5.1	101	2	Q9B863_SOROR	Q9b863_sorex ornat
1047	6	5.1	92	2	Q7MC41_VIBVY	Q7mc41_vibrio vuln	1120	6	5.1	101	2	Q9B867_SORMO	Q9b867_sorex monti
1048	6	5.1	92	2	Q9WBE3_9HIV1	Q9wbe3_human immun	1121	6	5.1	101	2	Q9B866_SORVA	Q9b866_sorex vagra
1049	6	5.1	93	2	Q5MIW1_AEDAL	Q5miw1_aedes albop	1122	6	5.1	101	2	Q9B0P4_SOROR	Q9b0p4_sorex ornat
1050	6	5.1	93	2	Q564M1_PANPO	Q564m1_panthera pa	1123	6	5.1	101	2	Q6K253_ORYSA	Q6k253_oryza sativ
1051	6	5.1	93	2	Q564M3_PANTA	Q564m3_panthera ti	1124	6	5.1	101	2	Q84C16_9ACTO	Q84c16_streptomyces
1052	6	5.1	93	2	Q564M4_PANTA	Q564m4_panthera ti	1125	6	5.1	101	2	Q4TL94_9SPHN	Q4tl94_erythroba
1053	6	5.1	93	2	Q564M0_PANPO	Q564m0_panthera pa	1126	6	5.1	101	2		

Q743k5_mycobacteri	1054	6	5.1	93	2	Q5BMW4_XANCA	Q5bmw4_xanthomonas
Q5zb39_oryza sativ	1055	6	5.1	93	2	Q5BMX2_XANCA	Q5bmX2_xanthomonas
Q6n83_oryza sativ	1056	6	5.1	93	2	Q8M240_FRANA	Q8m240_fraomyces nat
Q06779_mycobacteri	1057	6	5.1	93	2	Q7YGG0_9NEOB	Q7ygg0_bufo woodho
Q7u117_mycobacteri	1058	6	5.1	94	2	Q97ZB1_ECULI	Q97ze1_escherichia
Q820b9_anabaena sp	1059	6	5.1	94	2	Q9F576_ECULI	Q9f576_escherichia
Q98mz8_rhizobium l	1060	6	5.1	94	2	Q8FLI3_COREF	Q8fli3_corynebacte
Q98694_cuculus can	1061	6	5.1	94	2	Q62CJ8_BURMA	Q62cj8_burkholderi
Q99697_geococcyx c	1062	6	5.1	94	2	Q20499_9PASS	Q20499_andropadus
Q9hrw2_halobacteri	1063	6	5.1	94	2	Q20497_9PASS	Q20497_andropadus
Q69kp4_oryza sativ	1064	6	5.1	94	2	Q20501_9PASS	Q20501_andropadus
Q84qx9_oryza sativ	1065	6	5.1	94	2	Q9T6P3_9TELE	Q9t6p3_acrossochei
Q84kg8_mycobacteri	1066	6	5.1	94	2	Q20491_9PASS	Q20491_andropadus
Q7vvf5_bordetella	1067	6	5.1	94	2	Q20505_9PASS	Q20505_andropadus
Q7wb93_bordetella	1068	6	5.1	94	2	Q20489_9PASS	Q20489_andropadus
Q7wmr1_bordetella	1069	6	5.1	94	2	Q20509_9PASS	Q20509_andropadus
Q9a171_streptomyces	1070	6	5.1	94	2	Q20503_9PASS	Q20503_andropadus
Q9s2u4_streptomyces	1071	6	5.1	94	2	Q20495_9PASS	Q20495_andropadus
Q8p292_streptococ	1072	6	5.1	94	2	Q20493_9PASS	Q20493_andropadus
Q608f1_methylococcyx	1073	6	5.1	94	2	Q20511_9PASS	Q20511_phyllastrep
Q99696_dromococcyx	1074	6	5.1	94	2	Q21852_9PASS	Q21852_andropadus
Q99700_phaenocopha	1075	6	5.1	94	2	Q21782_9PASS	Q21782_andropadus
Q99703_tapera naev	1076	6	5.1	95	2	Q97516_9PRIM	Q97516_gorilla gor
Q99690_cercococcyx	1077	6	5.1	95	2	Q9T6P1_9TELE	Q9t6p1_acrossochei
Q99695_cuculus pol	1078	6	5.1	95	2	Q9T6N9_9TELE	Q9t6n9_acrossochei
Q99701_playa cayan	1079	6	5.1	95	2	Q20507_9PASS	Q20507_andropadus
Q99691_chrysococcy	1080	6	5.1	95	2	Q9T6P0_9TELE	Q9t6p0_acrossochei
Q874y3_chrysospora a	1081	6	5.1	95	2	Q9T6P2_9TELE	Q9t6p2

1127	6	5.1	101	2	Q5L271_GEOKA	Q5L271 geobacillus	1200	6	5.1	108	2	Q8IFJ3_9TRYP	Q8Ifj3 trypanosoma
1128	6	5.1	101	2	Q4THZ4_TETNG	Q4thz4 tetraodon n	1201	6	5.1	108	2	Q47625_CANLA	Q47625 canis latra
1129	6	5.1	101	2	Q9ZYL4_9GALL	Q9zyl4 callipepla	1202	6	5.1	108	2	Q8KGQ8_RHILO	Q8kgq8 rhizobium l
1130	6	5.1	101	2	Q9ZYL0_COLVI	Q9zyl0 colinus vir	1203	6	5.1	108	2	Q9ADQ3_STRCO	Q9adq3 streptomyc
1131	6	5.1	101	2	Q9ZVK6_0REPI	Q9zvk6 oreortyx pi	1204	6	5.1	108	2	Q9WYF2_THEMA	Q9wyf2 thermotoga
1132	6	5.1	101	2	Q9ZXP3_LORCA	Q9zxp3 lophortyx c	1205	6	5.1	108	2	Q9MPU4_PETRR	Q9mpu4 perissocoph
1133	6	5.1	101	2	Q9ZXUL_CASUS	Q9zxul callipepla	1206	6	5.1	108	2	Q9MPU7_PROAL	Q9mpu7 procnias a
1134	6	5.1	102	2	Q624Q0_CAEBR	Q624q0 caenorhabdi	1207	6	5.1	108	2	Q47626_CANLU	Q47626 canis lupus
1135	6	5.1	102	2	Q16462_CABEL	Q16462 caenorhabdi	1208	6	5.1	109	2	Q47622_CUOAL	Q47622 cuon alpinu
1136	6	5.1	102	2	Q44471_CABEL	Q44471 caenorhabdi	1209	6	5.1	109	2	Q8MPV6_9PASS	Q8mpv6 cotinga cay
1137	6	5.1	102	2	Q8UGT2_AGRF5	Q8ugt2 agrobacteri	1210	6	5.1	109	2	Q9MPW4_9PASS	Q9mpw4 ampelion ru
1138	6	5.1	102	2	Q83BAS_COXBU	Q83bas coxiella bu	1211	6	5.1	109	2	Q9DLB1_9HIV1	Q9dlb1 human immun
1139	6	5.1	102	2	Q9XP5_9HEPC	Q9xp5 hepatitis c	1212	6	5.1	109	2	Q47624_CHRBR	Q47624 chrysocyon
1140	6	5.1	102	2	Q7YQJ1_9TELE	Q7yqj1 laemonechima l	1213	6	5.1	110	2	Q47634_OTOME	Q47634 otocyon meg
1141	6	5.1	103	1	MGP_CHILA	Q6qn06 chinichilla	1214	6	5.1	110	2	Q47630_VULZE	Q47630 vulpes zerd
1142	6	5.1	103	1	MGP_PIG	Q8mj39 sus scrofa	1215	6	5.1	110	2	Q6ELF8_UCUSA	Q6elf8 cucumis sat
1143	6	5.1	103	2	Q75CL9_ASHGO	Q75cl9 ashyba goss	1216	6	5.1	110	2	Q7TWZ6_MYCBO	Q7twz6 mycobacteri
1144	6	5.1	103	2	Q47863_HIPNI	Q47863 hippotragus	1217	6	5.1	110	2	Q05850_MYCTU	Q05850 mycobacteri
1145	6	5.1	103	2	Q47864_ORYLE	Q47864 oryx leucor	1218	6	5.1	110	2	Q9WK01_9RETR	Q9wk01 rous sarcom
1146	6	5.1	103	2	Q7J5D2_HIPNI	Q7j5d2 hippotragus	1219	6	5.1	110	2	Q6THP1_9HIV1	Q6thp1 human immun
1147	6	5.1	103	2	Q7TM02_9ZZZZ	Q7tm02 uncultured	1220	6	5.1	110	2	Q47623_CANAU	Q47623 canis aureu
1148	6	5.1	103	2	Q5PU58_9TELE	Q5pu58 xyrauchen t	1221	6	5.1	111	2	Q47628_CANME	Q47628 canis mesom
1149	6	5.1	103	2	Q5PU56_9TELE	Q5pu56 xyrauchen t	1222	6	5.1	111	2	Q47629_DUSTH	Q47629 dusicyon th
1150	6	5.1	103	2	Q5PU54_9TELE	Q5pu54 xyrauchen t	1223	6	5.1	111	2	Q47637_PSEGT	Q47637 pseudalopec
1151	6	5.1	103	2	Q5PU53_9TELE	Q5pu53 xyrauchen t	1224	6	5.1	111	2	Q47632_PSEVE	Q47632 pseudalopec
1152	6	5.1	103	2	Q5PU52_9TELE	Q5pu52 xyrauchen t	1225	6	5.1	111	2	Q47635_PSEPC	Q47635 pseudalopec
1153	6	5.1	103	2	Q5PU50_9TELE	Q5pu50 xyrauchen t	1226	6	5.1	111	2	Q47641_VULMA	Q47641 vulpes macr
1154	6	5.1	103	2	Q5PU73_9TELE	Q5pu73 xyrauchen t	1227	6	5.1	111	2	Q47621_CANAD	Q47621 canis adust
1155	6	5.1	103	2	Q5PU76_9TELE	Q5pu76 xyrauchen t	1228	6	5.1	111	2	Q47627_CANME	Q47627 canis mesom
1156	6	5.1	104	2	Q8PVK3_METNA	Q8pvk3 methanosarc	1229	6	5.1	111	2	Q47636_PSEGR	Q47636 pseudalopec
1157	6	5.1	104	2	Q9A5N0_CAUCR	Q9asn0 caulobacter	1230	6	5.1	111	2	Q47639_PSEVE	Q47639 speothos ve
1158	6	5.1	104	2	Q6A094_METCA	Q6a094 methylococc	1231	6	5.1	111	2	Q47620_ATEMI	Q47620 atelocynus
1159	6	5.1	104	2	Q7YXG1_9NEOB	Q7yxg1 bufo woodho	1232	6	5.1	111	2	Q47642_VULVU	Q47642 vulpes vulp
1160	6	5.1	104	2	Q6TA18_9AVES	Q6ta18 coragyps at	1233	6	5.1	111	2	Q47631_LYCPI	Q47631 lycosa pict
1161	6	5.1	104	2	Q8WU5_9HIV1	Q8wu5 human immun	1234	6	5.1	111	2	Q34255_CANSI	Q34255 canis simen
1162	6	5.1	105	1	RL44Q_CANMA	P27074 candida mal	1235	6	5.1	111	2	Q5UUM0_MYRLE	Q5uum0 myrmothelul
1163	6	5.1	105	1	RL44Q_CANTR	P27075 candida tro	1236	6	5.1	111	2	Q9ENE3_9HIV1	Q9ene3 human immun
1164	6	5.1	105	2	Q9T268_9SCOM	Q9t268 acanthocybi	1237	6	5.1	111	2	THL_CLODI	P45362 clostridium
1165	6	5.1	105	2	Q9T267_9SCOM	Q9t267 acanthocybi	1238	6	5.1	112	1	Q70ZB5_CANFA	Q70zb5 canis fami
1166	6	5.1	106	1	INS2_KENLA	P12707 xenopus lae	1239	6	5.1	112	2	Q70ZC1_CANFA	Q70zc1 canis fami
1167	6	5.1	106	2	Q8X204_TALEM	Q8x204 talaromyces	1240	6	5.1	112	2	Q70ZC4_CANFA	Q70zc4 canis fami
1168	6	5.1	106	2	Q47640_UROCI	Q47640 urocyon cin	1241	6	5.1	112	2	Q4KWF4_9CHIR	Q4kwf4 myotis nipa
1169	6	5.1	106	2	Q47633_NYCPR	Q47633 nyctereutes	1242	6	5.1	112	2	Q4KWF5_9CHIR	Q4kwf5 myotis nipa
1170	6	5.1	106	2	Q5WHU8_BACSK	Q5whu8 bacillus cl	1243	6	5.1	112	2	Q4KWF6_MYOMS	Q4kwf6 myotis myet
1171	6	5.1	106	2	Q92LD2_RHIME	Q92ld2 rhizobium m	1244	6	5.1	112	2	Q4KWF7_9CHIR	Q4kwf7 myotis aura
1172	6	5.1	106	2	Q79297_TAUDA	Q79297 tauraco har	1245	6	5.1	112	2	Q4KWF9_9CHIR	Q4kwf9 myotis aura
1173	6	5.1	106	2	Q5QAY1_9HIV1	Q5qay1 human immun	1246	6	5.1	112	2	Q4KWF9_9CHIR	Q4kwf9 myotis aura
1174	6	5.1	107	2	Q6BYP5_DEBHA	Q6byp5 debaryomyce	1247	6	5.1	112	2	Q4KWF9_9CHIR	Q4kwf9 myotis aura
1175	6	5.1	107	2	Q47638_PSEBU	Q47638 pseudalopec	1248	6	5.1	112	2	Q93QG2_9MICO	Q93qg2 myotis aura
1176	6	5.1	107	2	Q5ULZ5_FRAAN	Q5ulz5 fragaria an	1249	6	5.1	112	2	Q9RV16_DEIRA	Q9rv16 deinococcus
1177	6	5.1	107	2	Q9LE17_HORVU	Q9le17 hordeum vul	1250	6	5.1	112	2	Q9NPT2_PIPPA	Q9npt2 pipra fasci
1178	6	5.1	107	2	Q4W4C8_HORVU	Q4w4c8 hordeum vul	1251	6	5.1	112	2	Q9NPT3_PIPCL	Q9npt3 piprites ch
1179	6	5.1	107	2	Q6AE31_LEIXX	Q6ae31 leifsonia x	1252	6	5.1	112	2	Q9NPT0_9PASS	Q9npt0 machaeropte
1180	6	5.1	107	2	Q72VG3_LEPIC	Q72vg3 leptospira	1253	6	5.1	112	2	Q9NPS9_9PASS	Q9nps9 machaeropte
1181	6	5.1	107	2	Q63KC5_BURPS	Q63kc5 burkholderi	1254	6	5.1	112	2	Q9NPS8_9PASS	Q9nps8 machaeropte
1182	6	5.1	107	2	Q89RS9_BRAJA	Q89rs9 bradyrhizob	1255	6	5.1	112	2	Q20983_9PSIT	Q20983 calyptrorhyn
1183	6	5.1	107	2	Q79285_9PASS	Q79285 thamnophilu	1256	6	5.1	112	2	Q73676_GEOOC	Q73676 geopsittacu
1184	6	5.1	107	2	Q79291_9CORV	Q79291 elminia lon	1257	6	5.1	112	2	Q73674_9PSIT	Q73674 neophema pe
1185	6	5.1	107	2	Q79276_9AVES	Q79276 coccyzus am	1258	6	5.1	112	2	Q73673_PZWA	Q73673 pezoporu w
1186	6	5.1	107	2	Q79282_9FURN	Q79282 furnarius r	1259	6	5.1	112	2	Q73672_PLAIC	Q73672 platycercus
1187	6	5.1	107	2	Q79280_9DEND	Q79280 dendrocolap	1260	6	5.1	112	2	Q73671_POLAT	Q73671 polytelis a
1188	6	5.1	107	2	Q79292_FULAM	Q79292 fulica amer	1261	6	5.1	112	2	Q4XPM6_PLACH	Q4xpm6 plasmodium
1189	6	5.1	107	2	Q79283_9PASS	Q79283 rupicola ru	1262	6	5.1	113	2	Q5KL20_MYCPR	Q5kl20 mycobacteri
1190	6	5.1	107	2	Q79290_9AVES	Q79290 phoebeastria	1263	6	5.1	113	2	Q5KL122_MYCRH	Q5kl122 mycobacteri
1191	6	5.1	107	2	Q63641_9ITRA	Q63641 anaeretes f	1264	6	5.1	113	2	Q9G6D6_9AVES	Q9g6d6 crinifer pi
1192	6	5.1	107	2	Q79273_COLAU	Q79273 colaptes au	1265	6	5.1	113	2	Q6WFD8_9HIV1	Q6wfd8 human immun
1193	6	5.1	107	2	Q79281_9PASS	Q79281 forficarius	1266	6	5.1	113	2	Q91WD8_9HIV1	Q91wd8 human immun
1194	6	5.1	107	2	Q79278_GEOCA	Q79278 geococcyx c	1267	6	5.1	113	2	Q91WD9_9HIV1	Q91wd9 human immun
1195	6	5.1	107	2	Q79296_STUVU	Q79296 sturnus vul	1268	6	5.1	113	2	Q96779_HABIN	Q96779 haemophilus
1196	6	5.1	107	2	Q79284_SAYPH	Q79284 sayornis ph	1269	6	5.1	114	2	Q5NWX4_AZOSE	Q5nwx4 azoarcus sp
1197	6	5.1	107	2	Q75972_9HIV1	Q75972 human immun	1270	6	5.1	114	2	Q82VJ9_NITEU	Q82vj9 nitrosomona
1198	6	5.1	107	2	Q79510_9HIV1	Q79510 human immun	1271	6	5.1	114	2	Q9GB95_9CHAR	Q9gb95 irediparra
1199	6	5.1	108	1	Y1976_CAUCR	Q9a6v7 caulobacter	1272	6	5.1	114	2		

1273	6	5..1	114	2	Q9G6E2_9AVES	Q9G6E2 tauraco per
1274	6	5..1	114	2	Q9GB96_9CHAR	Q9GB96 macroparra
1275	6	5..1	114	2	Q6L7A5_9PERO	Q6L7A5 pentacaros
1276	6	5..1	114	2	Q6L7A8_9PERO	Q6L7A8 pseudopentaca
1277	6	5..1	114	2	Q6L7A6_9PERO	Q6L7A6 pentacaros
1278	6	5..1	114	2	O10741_9HIV1	O10741 human immunu
1279	6	5..1	114	2	O10748_9HIV1	O10748 human immunu
1280	6	5..1	114	2	Q71437_9HIV1	Q71437 human immunu
1281	6	5..1	114	2	Q8Q6N7_9HIV1	Q8Q6N7 human immunu
1282	6	5..1	115	2	Q51PQ8_MAGGR	Q51PQ8 magnaporthes
1283	6	5..1	115	2	Q5NE33_WHEAT	Q5NE33 triticum ea
1284	6	5..1	115	2	Q8RP93_STRPN	Q8RP93 streptococ
1285	6	5..1	115	2	Q68AN2_9BURK	Q68AN2 ralstonia m
1286	6	5..1	115	2	Q4H971_9DEIO	Q4H971 deinooccoccus
1287	6	5..1	115	2	Q4NZF6_9DELT	Q4NZF6 anaerococcus
1288	6	5..1	115	2	Q8DP51_STRFR6	Q8DP51 streptococ
1289	6	5..1	115	2	Q8KS22_RALSO	Q8KS22 ralstonia s
1290	6	5..1	115	2	Q37P04_STRPN	Q37P04 streptococ
1291	6	5..1	115	2	Q8RWY9_DEIRA	Q8RWY9 deinooccoccus
1292	6	5..1	115	2	Q9GBA1_ACTAF	Q9GBA1 actophilor
1293	6	5..1	115	2	Q9GBA8_HYDC	Q9GBA8 hydrophasia
1294	6	5..1	115	2	Q9GBA0_JACJC	Q9GBA0 jacana spina
1295	6	5..1	115	2	Q9GB98_9CHAR	Q9GB98 jacana spina
1296	6	5..1	115	2	Q9GB94_9CHAR	Q9GB94 rostratula
1297	6	5..1	115	2	Q9GB97_METIN	Q9GB97 metopidius
1298	6	5..1	115	2	Q3UUQ0_9PASS	Q3UUQ0 microceru
1299	6	5..1	115	2	Q39130_9HIV1	Q39130 human immu
1300	6	5..1	115	2	Q90463_9HIV1	Q90463 human immu
1301	6	5..1	115	2	Q8BCV3_9HIV1	Q8BCV3 human immu
1302	6	5..1	115	2	Q6WFC6_9HIV1	Q6WFC6 human immu
1303	6	5..1	115	2	Q7ZE43_9HIV1	Q7ZE43 human immu
1304	6	5..1	115	2	Q9E155_9HIV1	Q9E155 human immu
1305	6	5..1	116	2	O47817_VULMA	O47817 vulpee macr
1306	6	5..1	116	2	Q4NBQ4_9NICC	Q4NBQ4 arthrobacte
1307	6	5..1	116	2	Q4NP74_9DELT	Q4NP74 anaeromyxob
1308	6	5..1	116	2	Q7YGF9_9NEOB	Q7YGF9 bufo woodh
1309	6	5..1	116	2	P89762_9HIV1	P89762 human immu
1310	6	5..1	116	2	Q7ZD82_9HIV1	Q7ZD82 human immu
1311	6	5..1	116	2	Q9WNQ4_9HIV1	Q9WNQ4 human immu
1312	6	5..1	117	2	Q51YA5_MAGGR	Q51YA5 magnaporth
1313	6	5..1	117	2	Q9HF55_ASHGO	Q9HF55 ashbya gos
1314	6	5..1	117	2	Q5NE30_WHEAT	Q5NE30 triticum aec
1315	6	5..1	117	2	Q5D3P9_9BACT	Q5D3P9 uncultured
1316	6	5..1	117	2	Q5D405_9BACT	Q5D405 uncultured
1317	6	5..1	117	2	Q5D402_9BACT	Q5D402 uncultured
1318	6	5..1	117	2	Q5D3Z0_9BACT	Q5D3Z0 uncultured
1319	6	5..1	117	2	Q5D3Y2_9BACT	Q5D3Y2 uncultured
1320	6	5..1	117	2	Q5D3Y1_9BACT	Q5D3Y1 uncultured
1321	6	5..1	117	2	Q5D3X7_9BACT	Q5D3X7 uncultured
1322	6	5..1	117	2	Q5D3X5_9BACT	Q5D3X5 uncultured
1323	6	5..1	117	2	Q5D3Y8_9BACT	Q5D3Y8 uncultured
1324	6	5..1	117	2	Q92WZ6_RHIME	Q92WZ6 rhizobium m
1325	6	5..1	117	2	Q9MPU0_9PASS	Q9MPU0 iodopileura
1326	6	5..1	117	2	Q9MPU3_PPRSC	Q9MPU3 pyroderus s
1327	6	5..1	117	2	Q9MPV5_PORPO	Q9MPV5 porphyrolac
1328	6	5..1	117	2	Q6BCU9_9HIV1	Q6BCU9 human immu
1329	6	5..1	117	2	Q6WFJ0_9HIV1	Q6WFJ0 human immu
1330	6	5..1	118	2	Q976E0_SULTO	Q976E0 sulfolobus
1331	6	5..1	118	2	Q5VKK2_PLAFA	Q5VKK2 plasmodium
1332	6	5..1	118	2	Q5VKK5_PLAFA	Q5VKK5 plasmodium
1333	6	5..1	118	2	Q5WRH8_PLAFA	Q5WRH8 plasmodium
1334	6	5..1	118	2	Q6V9G4_PLAFA	Q6V9G4 plasmodium
1335	6	5..1	118	2	Q6V9G6_PLAFA	Q6V9G6 plasmodium
1336	6	5..1	118	2	Q6V9H1_PLAFA	Q6V9H1 plasmodium
1337	6	5..1	118	2	Q5D3R3_9BACT	Q5D3R3 uncultured
1338	6	5..1	118	2	Q5D3U0_9BACT	Q5D3U0 uncultured
1339	6	5..1	118	2	Q5UJA7_9BACT	Q5UJA7 uncultured
1340	6	5..1	118	2	Q5UJA6_9BACT	Q5UJA6 uncultured
1341	6	5..1	118	2	Q5UJA5_9BACT	Q5UJA5 uncultured
1342	6	5..1	118	2	Q5D4L9_9BACT	Q5D4L9 uncultured
1343	6	5..1	118	2	Q5D3R4_9BACT	Q5D3R4 uncultured
1344	6	5..1	118	2	Q5D3R2_9BACT	Q5D3R2 uncultured

1346	6	5.1	118	2	Q5D3R1_9BACT
1347	6	5.1	118	2	Q5D3Q2_9BACT
1348	6	5.1	118	2	Q5D3P8_9BACT
1349	6	5.1	118	2	Q5D4H2_9BACT
1350	6	5.1	118	2	Q6M8RA_BDRBA
1351	6	5.1	118	2	P89758_9HIV1
1352	6	5.1	118	2	Q7ZK12_9HIV1
1353	6	5.1	119	2	Q9YA02_AERPE
1354	6	5.1	119	2	Q5YV14_ANOGA
1355	6	5.1	119	2	Q9MG24_CALKU
1356	6	5.1	119	2	Q9MFP8_NYCPR
1357	6	5.1	119	2	Q9MG25_CALPI
1358	6	5.1	119	2	Q9WQ16_SAGOE
1359	6	5.1	119	2	Q63VZ6_BURPS
1360	6	5.1	119	2	Q87P25_VIBPA
1361	6	5.1	119	2	Q9G6D9_CORCR
1362	6	5.1	119	2	Q67B7_9DEND
1363	6	5.1	119	2	Q5UUG3_9PICI
1364	6	5.1	119	2	Q9WQ24_9PICI
1365	6	5.1	119	2	Q9MPQ4_9PASS
1366	6	5.1	119	2	Q7ZK17_9HIV1
1367	6	5.1	119	2	Q8Q5T8_9HIV1
1368	6	5.1	120	2	Q9S7M4_HYLM
1369	6	5.1	120	2	Q9S7M7_HYLM
1370	6	5.1	120	2	Q9S7M9_HYLAG
1371	6	5.1	120	2	Q9S7M3_HYLAG
1372	6	5.1	120	2	Q9S7X2_HYLM
1373	6	5.1	120	2	Q9S7M5_HYLM
1374	6	5.1	120	2	Q9S7M6_HYLM
1375	6	5.1	120	2	Q9S7M8_HYLAG
1376	6	5.1	120	2	Q9S7M2_HYLAG
1377	6	5.1	120	2	Q9S7X1_HYLM
1378	6	5.1	120	2	Q9S7M1_HYLM
1379	6	5.1	120	2	Q9QA1_HYLM
1380	6	5.1	120	2	Q9AP14_HYLM
1381	6	5.1	120	2	Q4LP50_9BURK
1382	6	5.1	120	2	Q63XD8_BURPS
1383	6	5.1	120	2	Q62N42_BURMA
1384	6	5.1	120	2	Q6N2F0_RHOPA
1385	6	5.1	120	2	Q4S8C3_TETNG
1386	6	5.1	120	2	Q9T837_9PASS
1387	6	5.1	120	2	Q5UUT0_9PASS
1388	6	5.1	120	2	Q5UUS8_9PASS
1389	6	5.1	120	2	Q5UUS6_9PASS
1390	6	5.1	120	2	Q5UUG4_9PICI
1391	6	5.1	120	2	Q98XA1_9HIV1
1392	6	5.1	121	2	Q8WVB2_HUMAN
1393	6	5.1	121	2	Q9B066_BPMB1
1394	6	5.1	121	2	Q8LY88_9FURN
1395	6	5.1	121	2	Q6WFE0_9HIV1
1396	6	5.1	122	2	Q5NBP8_ORYSA
1397	6	5.1	122	2	Q605P8_METCA
1398	6	5.1	122	2	Q8LY88_9FURN
1399	6	5.1	122	2	Q8LY98_9FURN
1400	6	5.1	122	2	Q8LY90_9FURN
1401	6	5.1	122	2	Q5UUS0_9PASS
1402	6	5.1	122	2	Q8LJ35_9FURN
1403	6	5.1	122	2	Q5MR47_9HIV1
1404	6	5.1	122	2	Q5QAW9_9HIV1
1405	6	5.1	122	2	Q6IUW4_9HIV1
1406	6	5.1	122	2	Q6IUU5_9HIV1
1407	6	5.1	122	2	Q6IUU8_9HIV1
1408	6	5.1	122	2	Q6IUU9_9HIV1
1409	6	5.1	122	2	Q6IUW0_9HIV1
1410	6	5.1	122	2	Q6IUW1_9HIV1
1411	6	5.1	122	2	Q6IUW2_9HIV1
1412	6	5.1	122	2	Q6IUW4_9HIV1
1413	6	5.1	122	2	Q6IUW5_9HIV1
1414	6	5.1	122	2	Q6IUW6_9HIV1
1415	6	5.1	122	2	Q7ZG11_9HIV1
1416	6	5.1	123	2	Q6V9G9_PLAFA
1417	6	5.1	123	2	Q9S708_CABEL
1418	6	5.1	123	2	Q7MW19_ALCB

Q5d3r1	uncultured
Q5d3q2	uncultured
Q5d3p8	uncultured
Q5d4h2	uncultured
Q6mr4	bdellovibrion
P9758	human immun
Q7zk12	human immu
Q9ya02	aeropyrum f
Q5ty14	anopheles c
Q9mg24	calithrix
Q3mf8	nycterautea
Q9mg25	callithrix
Q3mg16	seguinus o
Q83vz6	bukholderia
Q87vz2	vibrio parre
Q9g6d9	corythaesol
Q87bp7	xiphocollap
Q3ug93	selenidera
Q3ug92	salicaria c
Q9mpq4	human immu
Q7zk17	human immu
Q8gt58	human immu
Q557w4	hylobates n
Q557w7	hylobates n
Q557w9	hylobates z
Q557k3	hylobates z
Q557x2	hylobates n
Q557x5	hylobates n
Q557w5	hylobates n
Q557w6	hylobates z
Q557w1	hylobates z
Q557x1	hylobates z
Q557w1	hylobates n
Q34qa1	hylobates n
Q34pi4	hylobates n
Q4lp80	bukholderia
Q83xd8	bukholderia
Q6zn42	rhopodopseud
Q6n2f0	rhopodopseud
Q48sc3	tetradodon r
Q8t837	hypocnemis
Q5uut0	hylophylax
Q3su88	hylophylax
Q5uus6	hylophylax
Q5ugu4	selenidera
Q39ex1	human immu
Q8wbv2	homo sapien
Q30b66	mycobacteri
Q4ly88	glyphorhynch
Q6wfe0	human immu
Q5nbp8	oryza sativa
Q405f8	methylocoeco
Q4ly88	glyphorhynch
Q4ly98	glyphorhynch
Q4ly90	microcaecum
Q5uuq5	microcaecum
Q4luh35	glyphorhynch
Q5mh47	human immu
Q3ghaw9	human immu
Q6liuv4	human immu
Q6liuv5	human immu
Q6liuv8	human immu
Q6liuv9	human immu
Q6liuw1	human immu
Q6liuw2	human immu
Q6liuw4	human immu
Q6liuw5	human immu
Q6liuw6	human immu
Q7zvq1	plasmidom
Q6vg99	caenorhabdit
Q7wx19	alcaligenes
Q5y0w8	caenorhabdit

1419	6	5.1	123	2	Q504B1_BRARE	Q504b1 brachydanio	1492	6	5.1	124	2	Q67BP6_9DEND	Q67bp6 dendrocolap
1420	6	5.1	123	2	Q9T829_9PASS	Q9t829 dryomphila	1493	6	5.1	124	2	Q67BP3_9PASS	Q67bp3 procnias nu
1421	6	5.1	123	2	Q9T830_9PASS	Q9t830 dryomphila	1494	6	5.1	124	2	Q5UUS7_9PASS	Q5uus7 hylophylax
1422	6	5.1	123	2	Q9T835_9PASS	Q9t835 dryomphila	1495	6	5.1	124	2	Q5UUS5_9PASS	Q5uus5 hylophylax
1423	6	5.1	123	2	Q9T840_9PASS	Q9t840 hypocnemis	1496	6	5.1	124	2	Q5UUS3_9PASS	Q5uus3 hylophylax
1424	6	5.1	123	2	Q9T845_MYLE	Q9t845 myrmotherul	1497	6	5.1	124	2	Q5UUR9_9PASS	Q5uur9 hylophylax
1425	6	5.1	123	2	Q9T826_9PASS	Q9t826 formicivora	1498	6	5.1	124	2	Q5UUN8_MYRHA	Q5uun8 myrmotherul
1426	6	5.1	123	2	Q9T847_9PASS	Q9t847 terenura hu	1499	6	5.1	124	2	Q5UUN5_MYRHA	Q5uun5 myrmotherul
1427	6	5.1	123	2	Q9T844_MYRIO	Q9t844 myrmotherul	1500	6	5.1	124	2	Q5UUK9_MYRIO	Q5uuk9 myrmotherul
1428	6	5.1	123	2	Q9T834_9PASS	Q9t834 dryomphila	1501	6	5.1	124	2	Q5UUK7_MYRIO	Q5uuk7 myrmotherul
1429	6	5.1	123	2	Q9T839_9PASS	Q9t839 hypocnemis	1502	6	5.1	124	2	Q5UUL1_9PASS	Q5uul1 phlegopsis
1430	6	5.1	123	2	Q9T841_9PASS	Q9t841 hypocnemis	1503	6	5.1	124	2	Q5UUL7_9PASS	Q5uul7 phlegopsis
1431	6	5.1	123	2	Q9T848_9PASS	Q9t848 formicarius	1504	6	5.1	124	2	Q5UUS2_9PASS	Q5uus2 hylophylax
1432	6	5.1	123	2	Q8L1A0_9FURN	Q8l1a0 formicarius	1505	6	5.1	124	2	Q5UUS4_9PASS	Q5uus4 microcerul
1433	6	5.1	123	2	Q9MPW3_9PASS	Q9mpw3 doliorinis s	1506	6	5.1	124	2	Q5UUS0_9PASS	Q5uus0 hylophylax
1434	6	5.1	123	2	Q8LY92_9FURN	Q8ly92 glyphorynch	1507	6	5.1	124	2	Q5UUK4_MYRIO	Q5uuk4 myrmotherul
1435	6	5.1	123	2	Q9T833_9PASS	Q9t833 dryomphila	1508	6	5.1	124	2	Q5UUL0_MYRIO	Q5uul0 myrmotherul
1436	6	5.1	123	2	Q9T843_MYRHA	Q9t843 myrmotherul	1509	6	5.1	124	2	Q5UUK6_MYRIO	Q5uuk6 myrmotherul
1437	6	5.1	123	2	Q9T832_9PASS	Q9t832 dryomphila	1510	6	5.1	124	2	Q5UUP1_MYRHA	Q5uup1 myrmotherul
1438	6	5.1	123	2	Q9T828_9DEND	Q9t828 lepidocolap	1511	6	5.1	124	2	Q5UUN9_MYRHA	Q5uun9 myrmotherul
1439	6	5.1	123	2	Q9T831_9PASS	Q9t831 dryomphila	1512	6	5.1	124	2	Q5UUM1_MYRLE	Q5uum1 myrmotherul
1440	6	5.1	123	2	Q8LY87_9FURN	Q8ly87 glyphorynch	1513	6	5.1	124	2	Q5UUM4_MYRLE	Q5uum4 myrmotherul
1441	6	5.1	123	2	Q9T838_9PASS	Q9t838 hypocnemis	1514	6	5.1	124	2	Q5UUS5_9PASS	Q5uus5 formicarius
1442	6	5.1	123	2	Q9T842_9PASS	Q9t842 hypocnemis	1515	6	5.1	124	2	Q9MPW1_9PASS	Q9mpw1 rupicola ru
1443	6	5.1	123	2	Q9T836_9PASS	Q9t836 formicivora	1516	6	5.1	124	2	Q9MPT1_9PASS	Q9mpt1 xenopipo at
1444	6	5.1	123	2	Q9T849_9PASS	Q9t849 pipra pipra	1517	6	5.1	124	2	Q5QAX7_9HIV1	Q5qax7 human immun
1445	6	5.1	123	2	Q9T827_9FURN	Q9t827 synallaxis	1518	6	5.1	124	2	Q5QAY0_9HIV1	Q5qay0 human immun
1446	6	5.1	123	2	Q5UUG1_9PIC1	Q5uug1 selenidera	1519	6	5.1	124	2	Q6QAZ9_9HIV1	Q6qaz9 human immun
1447	6	5.1	123	2	Q5UUG0_9PIC1	Q5uug0 selenidera	1520	6	5.1	124	2	Q6WFD7_9HIV1	Q6wfd7 human immun
1448	6	5.1	123	2	Q9T442_9PASS	Q9t442 hypocnemis	1521	6	5.1	124	2	Q6WFD9_9HIV1	Q6wfd9 human immun
1449	6	5.1	123	2	Q9T416_9PASS	Q9t416 dryomphila	1522	6	5.1	124	2	Q6WFE1_9HIV1	Q6wfe1 human immun
1450	6	5.1	123	2	Q8LY52_9FURN	Q8ly52 glyphorynch	1523	6	5.1	124	2	Q6WFE6_9HIV1	Q6wfe6 human immun
1451	6	5.1	123	2	Q8LUN8_9FURN	Q8lun8 glyphorynch	1524	6	5.1	124	2	Q6WFG8_9HIV1	Q6wfg8 human immun
1452	6	5.1	123	2	Q8LUN7_9FURN	Q8lun7 glyphorynch	1525	6	5.1	124	2	Q7ZGD6_9HIV1	Q7zgd6 human immun
1453	6	5.1	123	2	Q9T4P4_9PASS	Q9t4p4 dryomphila	1526	6	5.1	124	2	Q995U3_9HIV1	Q995u3 human immun
1454	6	5.1	123	2	Q9T3L3_9PASS	Q9t3l3 dryomphila	1527	6	5.1	124	2	Q7NRQ9_CHRVO	Q7nrq9 chromobacte
1455	6	5.1	123	2	Q8LYU6_9FURN	Q8lyu6 glyphorynch	1528	6	5.1	125	2	Q68Y58_ORYLA	Q68y58 oryzias lat
1456	6	5.1	123	2	Q6QTW8_9HIV1	Q6qtw8 human immun	1529	6	5.1	125	2	Q4TBF3_TETNG	Q4tbf3 tetraodon n
1457	6	5.1	123	2	Q6QTX9_9HIV1	Q6qtx9 human immun	1530	6	5.1	125	2	Q9MPQ5_9PASS	Q9mpq5 grallaria n
1458	6	5.1	123	2	Q6WFD3_9HIV1	Q6wfd3 human immun	1531	6	5.1	125	2	Q9MPQ7_GRARI	Q9mpq7 grallaria r
1459	6	5.1	123	2	Q6WFE7_9HIV1	Q6wfe7 human immun	1532	6	5.1	125	2	Q8LY99_9FURN	Q8ly99 glyphorynch
1460	6	5.1	123	2	Q6WFG9_9HIV1	Q6wfg9 human immun	1533	6	5.1	125	2	Q8LY94_9FURN	Q8ly94 glyphorynch
1461	6	5.1	123	2	Q7ZJY8_9HIV1	Q7zjy8 human immun	1534	6	5.1	125	2	Q8LY93_9FURN	Q8ly93 glyphorynch
1462	6	5.1	123	2	Q7ZJY9_9HIV1	Q7zjy9 human immun	1535	6	5.1	125	2	Q8LY82_9FURN	Q8ly82 glyphorynch
1463	6	5.1	124	2	Q9YBWB_AERPE	Q9ybwb aeropyrum p	1536	6	5.1	125	2	Q8LY86_9FURN	Q8ly86 glyphorynch
1464	6	5.1	124	2	Q6V9H2_PLAPA	Q6v9h2 plasmodium	1537	6	5.1	125	2	Q8LY95_9FURN	Q8ly95 glyphorynch
1465	6	5.1	124	2	Q8YNA1_ANASP	Q8yna1 anabaena sp	1538	6	5.1	125	2	Q8LY89_9FURN	Q8ly89 glyphorynch
1466	6	5.1	124	2	Q8C8I4_MOUSE	Q8c8i4 mus musculus	1539	6	5.1	125	2	Q8LY81_9FURN	Q8ly81 glyphorynch
1467	6	5.1	124	2	Q6DGP6_BRARE	Q6dgp6 brachydanio	1540	6	5.1	125	2	Q8LY97_9FURN	Q8ly97 glyphorynch
1468	6	5.1	124	2	Q9MPU5_9PASS	Q9mpu5 cephalopter	1541	6	5.1	125	2	Q8LY88_9FURN	Q8ly88 glyphorynch
1469	6	5.1	124	2	Q9MPW2_9PASS	Q9mpw2 rupicola pe	1542	6	5.1	125	2	Q8LY96_9FURN	Q8ly96 glyphorynch
1470	6	5.1	124	2	Q9MPW0_9PASS	Q9mpw0 phoenicircu	1543	6	5.1	125	2	Q9MPQ6_GRARI	Q9mpq6 grallaria r
1471	6	5.1	124	2	Q9MPT7_9PASS	Q9mpt7 schiiformis	1544	6	5.1	125	2	Q7YI41_9AVES	Q7yi41 crypturellu
1472	6	5.1	124	2	Q9MPV0_9PASS	Q9mpv0 lipaugus fu	1545	6	5.1	125	2	Q7YI38_9TURD	Q7yi38 turdus leuc
1473	6	5.1	124	2	Q9MPV4_9PASS	Q9mpv4 conioptilon	1546	6	5.1	125	2	Q67BS8_9PASS	Q67bs8 chamaeza ca
1474	6	5.1	124	2	Q9MPV2_9PASS	Q9mpv2 xipholena p	1547	6	5.1	125	2	Q67BS7_9PASS	Q67bs7 chamaeza mo
1475	6	5.1	124	2	Q9MPV6_LANEL	Q9mpv6 laniisoma e	1548	6	5.1	125	2	Q67BS5_9PASS	Q67bs5 grallariul
1476	6	5.1	124	2	Q9MPU9_9PASS	Q9mpu9 turdampelis	1549	6	5.1	125	2	Q67BR8_9PASS	Q67br8 grallaria e
1477	6	5.1	124	2	Q8LY80_9FURN	Q8ly80 glyphorynch	1550	6	5.1	125	2	Q67BR4_9PASS	Q67br4 formicaria
1478	6	5.1	124	2	Q8LY91_9FURN	Q8ly91 glyphorynch	1551	6	5.1	125	2	Q67BR3_9PASS	Q67br3 formicarius
1479	6	5.1	124	2	Q9MPV7_9PASS	Q9mpv7 ampeloides	1552	6	5.1	125	2	Q67BR1_9PASS	Q67br1 hylopesus b
1480	6	5.1	124	2	Q9MPV1_9PASS	Q9mpv1 haematoderu	1553	6	5.1	125	2	Q67BR0_9PASS	Q67br0 conopopaha
1481	6	5.1	124	2	Q9MPU1_9PASS	Q9mpu1 gymmoderu	1554	6	5.1	125	2	Q67BQ9_9PASS	Q67bq9 conopopaha
1482	6	5.1	124	2	Q9MPT9_9TYRA	Q9mpt9 pachyramphu	1555	6	5.1	125	2	Q67BQ8_9PASS	Q67bq8 pittasoma r
1483	6	5.1	124	2	Q9MPV3_9PASS	Q9mpv3 carpodectes	1556	6	5.1	125	2	Q67BQ7_9PASS	Q67bq7 pittasoma m
1484	6	5.1	124	2	Q9MPV9_PIPAR	Q9mpv9 pipreola ar	1557	6	5.1	125	2	Q67BQ6_9PASS	Q67bq6 lioseceles t
1485	6	5.1	124	2	Q9MPT8_9TYRA	Q9mpt8 pachyramphu	1558	6	5.1	125	2	Q67BQ3_9PASS	Q67bq3 rhinocrypta
1486	6	5.1	124	2	Q9MPV8_PIPCH	Q9mpv8 pipreola ch	1559	6	5.1	125	2	Q67BQ2_9PASS	Q67bq2 phlegopsis
1487	6	5.1	124	2	Q9MPT5_9TYRA	Q9mpt5 tityra caya	1560	6	5.1	125	2	Q67BQ1_9FURN	Q67bq1 furnarius r
1488	6	5.1	124	2	Q9MPT4_9TYRA	Q9mpt4 tityra inqu	1561	6	5.1	125	2		
1489	6	5.1	124	2	Q9MPU2_9PASS	Q9mpu2 querula pur	1562	6	5.1	125	2		
1490	6	5.1	124	2	Q7YGF6_9NEOB	Q7ygf6 bufo woodho	1563	6	5.1	125	2		
1491	6	5.1	124	2	Q67BP8_9FURN	Q67bp8 xenops minu	1564	6	5.1	125	2		

1565	6	5.1	125	2	Q67B01_9PASS	Q67bq1 thamnophilu	1638	6	5.1	127	2	Q6WFC2_9HIV1	Q6wfc2 human immun
1566	6	5.1	125	2	Q67B00_9TVRA	Q67bq0 corythopis	1639	6	5.1	127	2	Q6WFC9_9HIV1	Q6wfc9 human immun
1567	6	5.1	125	2	Q67B99_9PASS	Q67bp9 myrmormonis t	1640	6	5.1	127	2	Q6WFD1_9HIV1	Q6wfd1 human immun
1568	6	5.1	125	2	Q7Y7B8_9PASS	Q7Y7b8 formicivora	1641	6	5.1	127	2	Q6WFD2_9HIV1	Q6wfd2 human immun
1569	6	5.1	125	2	Q7Y670_9TVRA	Q7Y670 elaeenia chi	1642	6	5.1	127	2	Q6WFD6_9HIV1	Q6wfd6 human immun
1570	6	5.1	125	2	Q7Y7A7_9EMBE	Q7Y7a7 cypsnagra h	1643	6	5.1	127	2	Q6WFE4_9HIV1	Q6wfe4 human immun
1571	6	5.1	125	2	Q8LV79_9FURN	Q8lv79 glyphorynch	1644	6	5.1	127	2	Q6WFG7_9HIV1	Q6wfg7 human immun
1572	6	5.1	125	2	Q8LV78_9FURN	Q8lv78 glyphorynch	1645	6	5.1	127	2	Q6WFH0_9HIV1	Q6wfho human immun
1573	6	5.1	125	2	Q8LY84_9FURN	Q8ly84 glyphorynch	1646	6	5.1	127	2	Q6WFH1_9HIV1	Q6wfhi human immun
1574	6	5.1	125	2	Q8LV63_9UPMC	Q8lv63 eupetomema	1647	6	5.1	127	2	Q6WFI3_9HIV1	Q6wfi3 human immun
1575	6	5.1	125	2	Q8LV60_9FURN	Q8lv60 synallaxis	1648	6	5.1	127	2	Q6WFI5_9HIV1	Q6wfi5 human immun
1576	6	5.1	125	2	Q7Y7M2_9EMBE	Q7Y7m2 tangara cay	1649	6	5.1	127	2	Q6WFI6_9HIV1	Q6wfi6 human immun
1577	6	5.1	125	2	Q8LV49_9FURN	Q8lv49 glyphorynch	1650	6	5.1	127	2	Q6WFI7_9HIV1	Q6wfi7 human immun
1578	6	5.1	125	2	Q8LUP2_9FURN	Q8lup2 glyphorynch	1651	6	5.1	127	2	Q6WFI8_9HIV1	Q6wfi8 human immun
1579	6	5.1	125	2	Q8LUP2_9FURN	Q8lup2 glyphorynch	1652	6	5.1	127	2	Q6WFJ1_9HIV1	Q6wfj1 human immun
1580	6	5.1	125	2	Q8LUP1_9FURN	Q8lup1 glyphorynch	1653	6	5.1	127	2	Q6WFJ4_9HIV1	Q6wfj4 human immun
1581	6	5.1	125	2	Q8LUP0_9FURN	Q8lup0 glyphorynch	1654	6	5.1	127	2	Q6WFJ5_9HIV1	Q6wfj5 human immun
1582	6	5.1	125	2	Q8LUN9_9FURN	Q8lun9 glyphorynch	1655	6	5.1	127	2	Q6WFJ7_9HIV1	Q6wfj7 human immun
1583	6	5.1	125	2	Q7Y614_9EMBE	Q7y614 sporophila	1656	6	5.1	127	2	Q6WFJ9_9HIV1	Q6wfj9 human immun
1584	6	5.1	125	2	Q8LV14_9FURN	Q8lv14 glyphorynch	1657	6	5.1	127	2	Q6WFK1_9HIV1	Q6wfk1 human immun
1585	6	5.1	125	2	Q7Y6R8_9TURD	Q7y6r8 turdus leuc	1658	6	5.1	127	2	Q6WFK2_9HIV1	Q6wfk2 human immun
1586	6	5.1	125	2	Q7Y7H7_9EMBE	Q7Y7h7 tangara cay	1659	6	5.1	127	2	Q6WFK6_9HIV1	Q6wfk6 human immun
1587	6	5.1	125	2	Q7Y7V9_9DEND	Q7Y7v9 lepidocolap	1660	6	5.1	128	2	Q5KX96_GEOKA	Q5kx96 Geobacillus
1588	6	5.1	125	2	Q8LUV7_9FURN	Q8luy7 glyphorynch	1661	6	5.1	128	2	Q99IX0_9ZZZZ	Q99ix0 uncultured
1589	6	5.1	125	2	Q8LY85_9FURN	Q8ly85 glyphorynch	1662	6	5.1	128	2	Q6NTX9_XENLA	Q6ntx9 xenopus lae
1590	6	5.1	125	2	Q9MPR0_9PASS	Q9mpr0 grallaria r	1663	6	5.1	128	2	Q6DJR3_XENTR	Q6djr3 xenopus tro
1591	6	5.1	125	2	Q9MPQ2_9PASS	Q9mpq2 myrmochera	1664	6	5.1	128	2	Q6WFC5_9HIV1	Q6wfc5 human immun
1592	6	5.1	125	2	Q9MPQ3_9PASS	Q9mpq3 grallaricul	1665	6	5.1	128	2	Q6WFD0_9HIV1	Q6wfd0 human immun
1593	6	5.1	125	2	Q9MPQ1_9PASS	Q9mpq1 hylopezus f	1666	6	5.1	128	2	Q6WFD4_9HIV1	Q6wfd4 human immun
1594	6	5.1	125	2	Q9MPR2_9PASS	Q9mpr2 grallaria r	1667	6	5.1	128	2	Q6WFD5_9HIV1	Q6wfd5 human immun
1595	6	5.1	125	2	Q9MPQ9_9PASS	Q9mpq9 grallaria w	1668	6	5.1	128	2	Q6WFB2_9HIV1	Q6wfb2 human immun
1596	6	5.1	125	2	Q9MPR1_9PASS	Q9mpr1 grallaria b	1669	6	5.1	128	2	Q6WFB3_9HIV1	Q6wfb3 human immun
1597	6	5.1	125	2	Q74821_9HIV1	Q74821 human immun	1670	6	5.1	128	2	Q6WFB8_9HIV1	Q6wfb8 human immun
1598	6	5.1	126	1	TCP4_HUMAN	P53999 homo sapien	1671	6	5.1	128	2	Q6WFB9_9HIV1	Q6wfb9 human immun
1599	6	5.1	126	2	Q9MG26_CALJA	Q9mg26 callithrix	1672	6	5.1	128	2	Q6WFF0_9HIV1	Q6wff0 human immun
1600	6	5.1	126	2	Q9MG20_CALHU	Q9mg20 callithrix	1673	6	5.1	128	2	Q6WFF1_9HIV1	Q6wff1 human immun
1601	6	5.1	126	2	Q9MG22_CALAR	Q9mg22 callithrix	1674	6	5.1	128	2	Q6WFF2_9HIV1	Q6wff2 human immun
1602	6	5.1	126	2	Q9MG14_SAI0E	Q9mg14 saimiri oer	1675	6	5.1	128	2	Q6WFF3_9HIV1	Q6wff3 human immun
1603	6	5.1	126	2	Q9MG21_9PRIM	Q9mg21 callithrix	1676	6	5.1	128	2	Q6WFF4_9HIV1	Q6wff4 human immun
1604	6	5.1	126	2	Q9MG23_CALAU	Q9mg23 callithrix	1677	6	5.1	128	2	Q6WFF5_9HIV1	Q6wff5 human immun
1605	6	5.1	126	2	Q9MG13_ALOSA	Q9mg13 alouatta sa	1678	6	5.1	128	2	Q6WFF6_9HIV1	Q6wff6 human immun
1606	6	5.1	126	2	Q9MG19_CBBPY	Q9mg19 cebuella py	1679	6	5.1	128	2	Q6WFF9_9HIV1	Q6wff9 human immun
1607	6	5.1	126	2	Q9MG18_CBBPY	Q9mg18 cebuella py	1680	6	5.1	128	2	Q6WFG0_9HIV1	Q6wfg0 human immun
1608	6	5.1	126	2	Q9MG15_CALGO	Q9mg15 callimico g	1681	6	5.1	128	2	Q6WFG1_9HIV1	Q6wfg1 human immun
1609	6	5.1	126	2	Q9MG17_LBORO	Q9mg17 leontopithe	1682	6	5.1	128	2	Q6WFG2_9HIV1	Q6wfg2 human immun
1610	6	5.1	126	2	Q8LXP8_OTITA	Q8lxp8 otis tarda	1683	6	5.1	128	2	Q6WFG3_9HIV1	Q6wfg3 human immun
1611	6	5.1	126	2	Q8LXP8_OTITA	Q8lxp8 otis tarda	1684	6	5.1	128	2	Q6WFG4_9HIV1	Q6wfg4 human immun
1612	6	5.1	126	2	Q7Y6K0_9PASS	Q7y6k0 timalia pil	1685	6	5.1	128	2	Q6WFG5_9HIV1	Q6wfg5 human immun
1613	6	5.1	126	2	Q6WFB2_9HIV1	Q6wfb2 human immun	1686	6	5.1	128	2	Q6WFG2_9HIV1	Q6wfh2 human immun
1614	6	5.1	126	2	Q6WFC7_9HIV1	Q6wfc7 human immun	1687	6	5.1	128	2	Q6WFG3_9HIV1	Q6wfh3 human immun
1615	6	5.1	126	2	Q6WFI4_9HIV1	Q6wfi4 human immun	1688	6	5.1	128	2	Q6WFG4_9HIV1	Q6wfh4 human immun
1616	6	5.1	126	2	Q7ZJX4_9HIV1	Q7zjx4 human immun	1689	6	5.1	128	2	Q6WFG5_9HIV1	Q6wfh5 human immun
1617	6	5.1	126	2	Q7ZJY1_9HIV1	Q7zjy1 human immun	1690	6	5.1	128	2	Q6WFG6_9HIV1	Q6wfh6 human immun
1618	6	5.1	126	2	Q9QMC0_9HIV1	Q9qmc0 human immun	1691	6	5.1	128	2	Q6WFG7_9HIV1	Q6wfh7 human immun
1619	6	5.1	126	2	Q9WZ99_9HIV1	Q9wz99 human immun	1692	6	5.1	128	2	Q6WFG8_9HIV1	Q6wfh8 human immun
1620	6	5.1	127	1	ATPZ_BACP3	P09354 bacillus ps	1693	6	5.1	128	2	Q6WFI0_9HIV1	Q6wfi0 human immun
1621	6	5.1	127	2	Q61BA2_HUMAN	Q61ba2 homo sapien	1694	6	5.1	128	2	Q6WFI2_9HIV1	Q6wfi2 human immun
1622	6	5.1	127	2	Q7QBG9_ANOGA	Q7qbg9 anopheles g	1695	6	5.1	128	2	Q6WFI9_9HIV1	Q6wfi9 human immun
1623	6	5.1	127	2	Q5R6D0_PONPY	Q5r6d0 pongo pygma	1696	6	5.1	128	2	Q6WFO2_9HIV1	Q6wfo2 human immun
1624	6	5.1	127	2	Q4R947_MACFA	Q4r947 macaca fasc	1697	6	5.1	128	2	Q6WFO3_9HIV1	Q6wfo3 human immun
1625	6	5.1	127	2	Q5XUI6_GEOKA	Q5xui6 geobacillus	1698	6	5.1	128	2	Q6WFO4_9HIV1	Q6wfo4 human immun
1626	6	5.1	127	2	Q9CQH6_MOUSE	Q9cqh6 m mus muscu	1699	6	5.1	128	2	Q6WFO5_9HIV1	Q6wfo5 human immun
1627	6	5.1	127	2	P89623_9HIV1	P89623 human immun	1700	6	5.1	128	2	Q6WFO6_9HIV1	Q6wfo6 human immun
1628	6	5.1	127	2	P89627_9HIV1	P89627 human immun	1701	6	5.1	128	2	Q6WFO7_9HIV1	Q6wfo7 human immun
1629	6	5.1	127	2	Q5QAX0_9HIV1	Q5qax0 human immun	1702	6	5.1	128	2	Q6WFO8_9HIV1	Q6wfo8 human immun
1630	6	5.1	127	2	Q6WFB3_9HIV1	Q6wfb3 human immun	1703	6	5.1	128	2	Q6WFO9_9HIV1	Q6wfo9 human immun
1631	6	5.1	127	2	Q6WFB4_9HIV1	Q6wfb4 human immun	1704	6	5.1	128	2	Q6WFO0_9HIV1	Q6wfo0 human immun
1632	6	5.1	127	2	Q6WFB5_9HIV1	Q6wfb5 human immun	1705	6	5.1	129	2	Q700Q2_PSEPU	Q700q2 pseudomonas
1633	6	5.1	127	2	Q6WFB7_9HIV1	Q6wfb7 human immun	1706	6	5.1	129	2	Q5QAX3_9HIV1	Q5qax3 human immun
1634	6	5.1	127	2	Q6WFB8_9HIV1	Q6wfb8 human immun	1707	6	5.1	129	2	Q5QAX6_9HIV1	Q5qax6 human immun
1635	6	5.1	127	2	Q6WFB9_9HIV1	Q6wfb9 human immun	1708	6	5.1	130	2	Q4LMN4_9EURK	Q4lmn4 burkholderi
1636	6	5.1	127	2	Q6WFC0_9HIV1	Q6wfc0 human immun	1709	6	5.1	130	2	Q4FSY9_9GAMM	Q4fay9 psychrobact
1637	6	5.1	127	2	Q6WFC1_9HIV1	Q6wfc1 human immun	1710	6	5.1	130	2	Q7TWM4_MYCBO	Q7twm4 mycobacteri

1711	6	5.1	130	2	050393 MYCTU	050393 mycobacteri	1784	6	5.1	133	2	Q7Y7P0_9PEZI	Q7Y7P0 exophiala c
1712	6	5.1	130	2	Q5Q196_IDILO	Q5Q196 idiomarina	1785	6	5.1	133	2	Q7Y795_9PEZI	Q7Y795 exophiala a
1713	6	5.1	130	2	Q5Q196_IDILO	Q5Q196 idiomarina	1786	6	5.1	133	2	Q7Y645_RHIAT	Q7Y645 rhinocladie
1714	6	5.1	132	2	Q956F3_9BASI	Q956F3 rhodotorula	1787	6	5.1	133	2	Q7Y813_9PEZI	Q7Y813 exophiala s
1715	6	5.1	132	2	Q956G2_9BASI	Q956G2 sporobolomy	1788	6	5.1	133	2	Q7Y6W1_9ASCO	Q7Y6W1 fonsecaea p
1716	6	5.1	132	2	Q956J2_9BASI	Q956J2 rhodotorula	1789	6	5.1	133	2	Q7Y6W2_9ASCO	Q7Y6W2 fonsecaea c
1717	6	5.1	132	2	Q956I3_9BASI	Q956I3 leucosporid	1790	6	5.1	133	2	Q7Y806_9PEZI	Q7Y806 exophiala m
1718	6	5.1	132	2	Q956J0_9BASI	Q956J0 rhodotorula	1791	6	5.1	133	2	Q7Y6F6_9PEZI	Q7Y6F6 aspergillus
1719	6	5.1	132	2	Q956G3_9BASI	Q956G3 sporidiobol	1792	6	5.1	133	2	Q7Y6F3_9PEZI	Q7Y6F3 emericella
1720	6	5.1	132	2	Q7R0P3_GIALA	Q7R0P3 giardia lam	1793	6	5.1	133	2	Q7Y7Y3_9PEZI	Q7Y7Y3 exophiala c
1721	6	5.1	132	2	Q8RNW2_RHIME	Q8RNW2 rhizobium m	1794	6	5.1	133	2	Q7Y7Y2_EXOJE	Q7Y7Y2 exophiala j
1722	6	5.1	132	2	Q9AP76_9BACT	Q9AP76 uncultured	1795	6	5.1	133	2	Q7Y7Y1_9PEZI	Q7Y7Y1 exophiala m
1723	6	5.1	132	2	Q5UUI5_9PASS	Q5UUI5 rhagmatorhi	1796	6	5.1	133	2	Q7Y6R7_9PEZI	Q7Y6R7 emericella
1724	6	5.1	132	2	Q5L101_9HIV1	Q5L101 human immun	1797	6	5.1	133	2	Q7Y6R6_9PEZI	Q7Y6R6 emericella
1725	6	5.1	133	2	Q8SUZ8_EXOJE	Q8SUZ8 exophiala j	1798	6	5.1	133	2	Q8SE93_ASPPA	Q8SE93 aspergillus
1726	6	5.1	133	2	Q9MR71_9PEZI	Q9MR71 aspergillus	1799	6	5.1	133	2	Q8SE95_ASPOR	Q8SE95 aspergillus
1727	6	5.1	133	2	Q8S7Z5_9PEZI	Q8S7Z5 exophiala b	1800	6	5.1	133	2	Q8SE97_ASPPL	Q8SE97 aspergillus
1728	6	5.1	133	2	Q8S7Z6_EXOJE	Q8S7Z6 exophiala j	1801	6	5.1	133	2	Q8SE92_9PEZI	Q8SE92 aspergillus
1729	6	5.1	133	2	Q8M6X9_ASPSO	Q8M6X9 aspergillus	1802	6	5.1	133	2	Q8SE94_ASPOR	Q8SE94 aspergillus
1730	6	5.1	133	2	Q9MR70_9PEZI	Q9MR70 aspergillus	1803	6	5.1	133	2	Q8LUL3_ASPNO	Q8LUL3 aspergillus
1731	6	5.1	133	2	Q7YGT8_9PEZI	Q7YGT8 aspergillus	1804	6	5.1	133	2	Q7Y6R5_ASPAW	Q7Y6R5 aspergillus
1732	6	5.1	133	2	Q7YGT7_9PEZI	Q7YGT7 aspergillus	1805	6	5.1	133	2	Q7Y6R4_9PEZI	Q7Y6R4 aspergillus
1733	6	5.1	133	2	Q7YGT6_9PEZI	Q7YGT6 aspergillus	1806	6	5.1	133	2	Q7Y6R3_ASPNG	Q7Y6R3 aspergillus
1734	6	5.1	133	2	Q7YGS6_9PEZI	Q7YGS6 aspergillus	1807	6	5.1	133	2	Q7Y6R2_9PEZI	Q7Y6R2 aspergillus
1735	6	5.1	133	2	Q7YGS5_ASPRE	Q7YGS5 aspergillus	1808	6	5.1	133	2	Q7Y6R0_ASPPT	Q7Y6R0 aspergillus
1736	6	5.1	133	2	Q7YGS4_9PEZI	Q7YGS4 eurotium ru	1809	6	5.1	133	2	Q69ME1_ORYSA	Q69ME1 oryzaeativ
1737	6	5.1	133	2	Q7YGR7_9PEZI	Q7YGR7 emericella	1810	6	5.1	133	2	Q69YF7_AZOSE	Q69YF7 azoarcus fl
1738	6	5.1	133	2	Q7YGR6_9PEZI	Q7YGR6 emericella	1811	6	5.1	133	2	Q6A2U3_APOFL	Q6A2U3 apodemus p
1739	6	5.1	133	2	Q7YGR5_9PEZI	Q7YGR5 emericella	1812	6	5.1	133	2	Q5IEY2_9FALC	Q5IEY2 spizaetus c
1740	6	5.1	133	2	Q7YGR4_9PEZI	Q7YGR4 emericella	1813	6	5.1	133	2	Q6W8R9_9HIV1	Q6W8R9 human immun
1741	6	5.1	133	2	Q7YGR3_9PEZI	Q7YGR3 emericella	1814	6	5.1	133	2	Q59G24_HUMAN	Q59G24 homo sapien
1742	6	5.1	133	2	Q7YGR2_9PEZI	Q7YGR2 emericella	1815	6	5.1	133	2	Q7D0A2_AGR5	Q7D0A2 agrobacteri
1743	6	5.1	133	2	Q7YGR1_9PEZI	Q7YGR1 emericella	1816	6	5.1	134	2	Q03806_9GOBI	Q03806 tridentiger
1744	6	5.1	133	2	Q7YGR0_9PEZI	Q7YGR0 emericella	1817	6	5.1	134	2	Q03803_9GOBI	Q03803 tridentiger
1745	6	5.1	133	2	Q7YEV4_TALEM	Q7YEV4 talaromyces	1818	6	5.1	134	2	Q03796_9GOBI	Q03796 tridentiger
1746	6	5.1	133	2	Q7YEV3_9PEZI	Q7YEV3 emericella	1819	6	5.1	134	2	Q03799_9GOBI	Q03799 tridentiger
1747	6	5.1	133	2	Q7YEV2_9PEZI	Q7YEV2 emericella	1820	6	5.1	134	2	Q03791_9GOBI	Q03791 tridentiger
1748	6	5.1	133	2	Q7YEV1_9PEZI	Q7YEV1 exophiala p	1821	6	5.1	134	2	Q03790_9GOBI	Q03790 tridentiger
1749	6	5.1	133	2	Q7YEX5_9PEZI	Q7YEX5 exophiala s	1822	6	5.1	134	2	Q03804_9GOBI	Q03804 tridentiger
1750	6	5.1	133	2	Q7YEX4_EXOJE	Q7YEX4 exophiala j	1823	6	5.1	134	2	Q03795_9GOBI	Q03795 tridentiger
1751	6	5.1	133	2	Q7YEX3_9PEZI	Q7YEX3 phialophora	1824	6	5.1	134	2	Q5QAX5_9HIV1	Q5QAX5 human immun
1752	6	5.1	133	2	Q7YEX2_9PEZI	Q7YEX2 exophiala d	1825	6	5.1	134	2	Q6IUC0_STRMT	Q6IUC0 streptococ
1753	6	5.1	133	2	Q7YEX1_9PEZI	Q7YEX1 rhinocladie	1826	6	5.1	134	2	Q8EJT0_STRPN	Q8EJT0 streptococ
1754	6	5.1	133	2	Q7YEX0_9PEZI	Q7YEX0 exophiala h	1827	6	5.1	135	2	Q8RTI2_9BACT	Q8RTI2 uncultured
1755	6	5.1	133	2	Q7YEW9_9PEZI	Q7YEW9 exophiala s	1828	6	5.1	135	2	Q9G3N8_9RODE	Q9G3N8 eospalax ba
1756	6	5.1	133	2	Q7YEW8_9PEZI	Q7YEW8 exophiala s	1829	6	5.1	135	2	Q9G0Q5_9RODE	Q9G0Q5 eospalax ba
1757	6	5.1	133	2	Q7YEW6_9PEZI	Q7YEW6 capronia ma	1830	6	5.1	135	2	Q5QAX9_9HIV1	Q5QAX9 human immun
1758	6	5.1	133	2	Q7YEW5_9PEZI	Q7YEW5 rhinocladie	1831	6	5.1	135	2	Q6B792_9HIV1	Q6B792 human immun
1759	6	5.1	133	2	Q7YEW4_9PEZI	Q7YEW4 capronia se	1832	6	5.1	135	2	Q8ABE7_9HIV1	Q8ABE7 human immun
1760	6	5.1	133	2	Q7YEW3_PHIVE	Q7YEW3 phialophora	1833	6	5.1	136	2	Q7YD66_RABIT	Q7YD66 cryctolagus
1761	6	5.1	133	2	Q7YEW2_EXOJE	Q7YEW2 exophiala j	1834	6	5.1	136	2	Q9QX4_9HIV1	Q9QX4 human immun
1762	6	5.1	133	2	Q7YEW1_9PEZI	Q7YEW1 exophiala a	1835	6	5.1	137	2	Q5QAX8_9HIV1	Q5QAX8 human immun
1763	6	5.1	133	2	Q7YEW0_9PEZI	Q7YEW0 exophiala m	1836	6	5.1	137	2	YCF81_NBPOL	YCF81 nephroselm
1764	6	5.1	133	2	Q7YEV9_9PEZI	Q7YEV9 exophiala l	1837	6	5.1	138	2	Q31061_MYXXA	Q31061 myxococcus
1765	6	5.1	133	2	Q7YEV8_EXOJE	Q7YEV8 exophiala f	1838	6	5.1	138	2	Q91WQ2_MOUSE	Q91WQ2 mus musculu
1766	6	5.1	133	2	Q7YEV7_9PEZI	Q7YEV7 cadophora f	1839	6	5.1	138	2	Q7ZG58_9HIV1	Q7ZG58 human immun
1767	6	5.1	133	2	Q7YEV6_9PEZI	Q7YEV6 exophiala s	1840	6	5.1	139	2	Q9J6D8_9HIV1	Q9J6D8 human immun
1768	6	5.1	133	2	Q7YEV5_9PEZI	Q7YEV5 exophiala p	1841	6	5.1	139	2	Q4TJR5_9SPFN	Q4TJR5 erythroba
1769	6	5.1	133	2	Q7YEV4_9PEZI	Q7YEV4 cladophialo	1842	6	5.1	139	2	Q7TYB4_MYCBO	Q7TYB4 mycobacteri
1770	6	5.1	133	2	Q7Y9B5_ASPFI	Q7Y9B5 aspergillus	1843	6	5.1	139	2	Q8XX84_RALSO	Q8XX84 ralestonia s
1771	6	5.1	133	2	Q7Y9B4_9PEZI	Q7Y9B4 aspergillus	1844	6	5.1	139	2	Q9JUQ5_NEIMA	Q9JUQ5 neisseria m
1772	6	5.1	133	2	Q7Y9B2_9PEZI	Q7Y9B2 aspergillus	1845	6	5.1	139	2	P95002_MYCTU	P95002 mycobacteri
1773	6	5.1	133	2	Q7Y9B1_9PEZI	Q7Y9B1 aspergillus	1846	6	5.1	140	2	CRCB2_NOCTU	CRCB2 nocardia fa
1774	6	5.1	133	2	Q7Y9B0_9PEZI	Q7Y9B0 eurotium to	1847	6	5.1	140	2	Q9Y8J9_HALVO	Q9Y8J9 halobacteri
1775	6	5.1	133	2	Q8S7Z9_ASPNG	Q8S7Z9 aspergillus	1848	6	5.1	140	2		
1776	6	5.1	133	2	Q8SEPT_EXOJE	Q8SEPT exophiala j	1849	6	5.1	140	2		
1777	6	5.1	133	2	Q7Y6M2_HORWE	Q7Y6M2 hortaea wer	1850	6	5.1	140	2		
1778	6	5.1	133	2	Q7Y700_9PEZI	Q7Y700 eurotium to	1851	6	5.1	140	2		
1779	6	5.1	133	2	Q8SE18_EXOJE	Q8SE18 aspergillus	1852	6	5.1	140	2		
1780	6	5.1	133	2	Q8SE17_EXOJE	Q8SE17 exophiala j	1853	6	5.1	140	2		
1781	6	5.1	133	2	Q8SE16_EXOJE	Q8SE16 exophiala j	1854	6	5.1	140	2		
1782	6	5.1	133	2	Q7Y7N9_EXODE	Q7Y7N9 exophiala d	1855	6	5.1	140	2		
1783	6	5.1	133	2	Q7Y7N8_9PEZI	Q7Y7N8 exophiala h	1856	6	5.1	140	2		

1857	6	5.1	140	2	Q86Q31_HYDAT	Q86q31 hydra atten	1930	6	5.1	148	2	Q63SB1_BURPS	Q63sbl burkholderi
1858	6	5.1	140	2	Q4NV76_9DELTA	Q4nv76 anaeromyxob	1931	6	5.1	148	2	Q62LR4_BURMA	Q62lr4 burkholderi
1859	6	5.1	140	2	Q4NXB0_9DELTA	Q4nxb0 anaeromyxob	1932	6	5.1	148	2	Q6CZ11_ERWCT	Q6cz11 erwinia car
1860	6	5.1	140	2	Q8RBN8_THETIN	Q8rbn8 thermoanaer	1933	6	5.1	149	1	CYB_CERN	P82047 cervus nipp
1861	6	5.1	140	2	Q73U97_MYCPA	Q73u97 mycobacteri	1934	6	5.1	149	2	Q9XNN2_CEREL	Q9xnn2 cervus elap
1862	6	5.1	140	2	Q7YGL9_9PASS	Q7ygl9 liocichla s	1935	6	5.1	149	2	Q7YD67_RABIT	Q7y67 oryctolagus
1863	6	5.1	141	2	Q5B115_EMENI	Q5b115 aspergillus	1936	6	5.1	149	2	Q7YD65_RABIT	Q7y65 oryctolagus
1864	6	5.1	141	2	Q5SMC4_CRYNE	Q5smc4 cryptococcus	1937	6	5.1	149	2	Q53B44_9BACT	Q53b44 uncultured
1865	6	5.1	141	2	Q5K8E9_CRYNE	Q5k8e9 cryptococcus	1938	6	5.1	149	2	Q53B38_9BACT	Q53b38 uncultured
1866	6	5.1	141	2	Q7YGN5_9SYLV	Q7ygn5 garriulax er	1939	6	5.1	149	2	Q53B27_METPL	Q53b27 methylocell
1867	6	5.1	141	2	Q7YGM0_9PASS	Q7ygm0 liocichla p	1940	6	5.1	149	2	Q7VMZ9_BORPE	Q7vmz9 bordetella
1868	6	5.1	142	1	YBR4_YEAST	P82420 saccharomyc	1941	6	5.1	149	2	Q7W7V2_BORPA	Q7w7v2 bordetella
1869	6	5.1	142	2	Q5AYF8_EMENI	Q5ayf8 aspergillus	1942	6	5.1	149	2	Q7WL92_BORBR	Q7wl92 bordetella
1870	6	5.1	142	2	Q75813_ASHGO	Q75813 ashbya goos	1943	6	5.1	149	2	Q87Q11_VIBPA	Q87q11 vibrio para
1871	6	5.1	142	2	Q4NYU0_9DELTA	Q4nyu0 anaeromyxob	1944	6	5.1	149	2	Q7V2W1_PROMP	Q7v2w1 prochloroco
1872	6	5.1	142	2	Q741K9_MYCPA	Q741k9 mycobacteri	1945	6	5.1	150	2	Q7YD54_RABIT	Q7y54 oryctolagus
1873	6	5.1	142	2	Q8A3B5_BACTN	Q8a3b5 bacteroides	1946	6	5.1	150	2	Q52662_RHOCA	Q52662 rhodobacter
1874	6	5.1	142	2	Q92DV3_LISIN	Q92dv3 listeria in	1947	6	5.1	150	2	Q911U6_PSEAE	Q911u6 pseudomonas
1875	6	5.1	142	2	Q8N557_RHOPA	Q8n557 rhodospseudo	1948	6	5.1	150	2	Q98FN4_RHIL0	Q98fn4 rhizobium l
1876	6	5.1	142	2	Q7VUV8_BORPE	Q7vuv8 bordetella	1949	6	5.1	150	2	Q98LW8_RHIL0	Q98lw8 rhizobium l
1877	6	5.1	142	2	Q6V160_9PERO	Q6v160 etheostoma	1950	6	5.1	150	2	Q65LF7_BACLD	Q65lf7 bacillus l
1878	6	5.1	142	2	Q6V156_9PERO	Q6v156 etheostoma	1951	6	5.1	150	2	Q6BCR6_9HIV1	Q6bcr6 human immun
1879	6	5.1	142	2	Q6V155_9PERO	Q6v155 etheostoma	1952	6	5.1	151	2	Q5V4C8_HALMA	Q5v4c8 haloarcula
1880	6	5.1	142	2	Q6V147_9PERO	Q6v147 etheostoma	1953	6	5.1	151	2	Q9YEG3_AERPE	Q9yeg3 aeropyrum p
1881	6	5.1	142	2	Q6V144_9PERO	Q6v144 etheostoma	1954	6	5.1	151	2	Q4PN60_IXOSC	Q4pn60 ixodes scap
1882	6	5.1	142	2	Q6V139_9PERO	Q6v139 etheostoma	1955	6	5.1	151	2	Q83SM6_SALTI	Q83sw6 salmonella
1883	6	5.1	142	2	Q6V165_9PERO	Q6v165 etheostoma	1956	6	5.1	151	2	Q57J05_SALCH	Q57j05 salmonella
1884	6	5.1	142	2	Q5Y2J3_9HIV1	Q5y2j3 human immun	1957	6	5.1	151	2	Q5PMS9_SALPA	Q5pms9 salmonella
1885	6	5.1	142	2	Q7ZJ15_9HIV1	Q7zj15 human immun	1958	6	5.1	151	2	Q63X36_BURPS	Q63x36 burkholderi
1886	6	5.1	143	2	Q8U0V1_PYRFU	Q8u0v1 pyrococcus	1959	6	5.1	151	2	Q8Z300_SALTI	Q8z300 salmonella
1887	6	5.1	143	2	Q8ACB5_LEIXX	Q8ace8 leifsonia x	1960	6	5.1	151	2	Q8ZM04_SALTY	Q8zm04 salmonella
1888	6	5.1	143	2	Q7NKY5_GLOVI	Q7nky5 gloeobacter	1961	6	5.1	151	2	Q6D8G3_ERWCT	Q6d8g3 erwinia car
1889	6	5.1	143	2	Q6R5F5_MOUSE	Q6r5f5 mus musculu	1962	6	5.1	151	2	Q9K7J5_BACHD	Q9k7j5 bacillus ha
1890	6	5.1	143	2	Q933H2_9GAMA	Q933h2 callitrichi	1963	6	5.1	151	2	Q9D3I5_MOUSE	Q9d3i5 mus musculu
1891	6	5.1	144	2	Q05980_SULAC	Q05980 sulfolobus	1964	6	5.1	151	2	Q91EN1_9PICO	Q91en1 foot-and-mo
1892	6	5.1	144	2	Q9HQ09_HALSA	Q9hq09 halobacteri	1965	6	5.1	151	2	Q4LDT2_RANJA	Q4ldt2 rana japoni
1893	6	5.1	144	2	Q51714_BORBU	Q51714 borrelia bu	1966	6	5.1	152	1	Y396_MYCE	P47636 mycoplasma
1894	6	5.1	144	2	Q8EFT3_SHEON	Q8eft3 shewanella	1967	6	5.1	152	2	Q75D17_ASHGO	Q75d17 ashbya goos
1895	6	5.1	144	2	Q9G5T7_AMPOC	Q9g5t7 amphiprion	1968	6	5.1	152	2	Q7YD58_RABIT	Q7y58 oryctolagus
1896	6	5.1	144	2	Q9G5T6_AMPOC	Q9g5t6 amphiprion	1969	6	5.1	152	2	Q05277_CHAYP	Q05277 chlorante-a
1897	6	5.1	144	2	Q7YGL6_9PASS	Q7ygl6 mnlia cyano	1970	6	5.1	152	2	Q4V066_XANCP	Q4v066 xanthomonas
1898	6	5.1	144	2	Q9G1V5_AMPOC	Q9g1v5 amphiprion	1971	6	5.1	152	2	Q6GYJ8_XANCT	Q6gyj8 uncultured
1899	6	5.1	144	2	Q70RA3_9HIV1	Q70ra3 human immun	1972	6	5.1	152	2	Q4PQX5_9GAMM	Q4pqx5 psychrobact
1900	6	5.1	145	1	CYB_ASPFL	P58629 aspergillus	1973	6	5.1	152	2	Q8PDX4_XANCP	Q8pdx4 xanthomonas
1901	6	5.1	145	2	Q7YCU7_9FULM	Q7ycu7 arion anthr	1974	6	5.1	152	2	Q9B190_LIZAU	Q9b190 liza aurata
1902	6	5.1	145	2	Q9TEF6_ODOHE	Q9tef6 odocolleus	1975	6	5.1	152	2	Q9B189_MUGCA	Q9b189 mugil capit
1903	6	5.1	145	2	Q6KEX0_METPL	Q6kex0 methylocell	1976	6	5.1	152	2	Q9B187_CHELBA	Q9b187 chelon labr
1904	6	5.1	145	2	Q5KX69_GROKA	Q5kx69 geobacillus	1977	6	5.1	153	1	CYB_CERN	P82045 cervus nipp
1905	6	5.1	145	2	Q9B8Y3_9PERC	Q9b8y9 liparia fab	1978	6	5.1	153	2	Q8N2F9_HUMAN	Q8n2f9 homo sapien
1906	6	5.1	145	2	Q8M4S3_9SYLV	Q8m4s3 stachyris s	1979	6	5.1	153	2	Q5C2N3_SCHJA	Q5c2n3 schistosoma
1907	6	5.1	146	2	Q50989_NEIGO	Q50989 neisseria g	1980	6	5.1	153	2	Q7YD70_RABIT	Q7y70 oryctolagus
1908	6	5.1	146	2	Q5F9W9_NEIGI	Q5f9w9 neisseria g	1981	6	5.1	153	2	Q9X691_HELHP	Q9x691 helicobacte
1909	6	5.1	146	2	Q6KEX7_METSI	Q6kex7 methylocell	1982	6	5.1	153	2	Q7UK00_RHOBA	Q7uk00 rhodopirell
1910	6	5.1	146	2	Q6KEX6_METTU	Q6kex6 methylocell	1983	6	5.1	153	2	Q7UW76_RHOBA	Q7uw76 rhodopirell
1911	6	5.1	146	2	Q6KEX1_METPL	Q6kex1 methylocell	1984	6	5.1	153	2	Q8XAX9_ECO57	Q8xax9 escherichia
1912	6	5.1	146	2	Q7P194_CHRVO	Q7p194 chromobact	1985	6	5.1	153	2	Q7Z5P6_DESVH	Q7z5p6 desulfovibr
1913	6	5.1	146	2	Q89K06_BRAJA	Q89k06 bradyrhizob	1986	6	5.1	154	1	CYB_CERNA	P82048 cervus nipp
1914	6	5.1	146	2	Q66542_AQUAE	Q66542 aquifex aeo	1987	6	5.1	154	1	CYB_CERNH	P82046 cervus nipp
1915	6	5.1	146	2	Q90BC8_9HIV1	Q90bc8 human immun	1988	6	5.1	154	1	CYB_CERNP	P82046 cervus nipp
1916	6	5.1	146	2	Q90BC9_9HIV1	Q90bc9 human immun	1989	6	5.1	154	2	Q9XM12_CEREL	Q9xm12 cervus elap
1917	6	5.1	146	2	Q90BD0_9HIV1	Q90bd0 human immun	1990	6	5.1	154	2	Q7YD68_RABIT	Q7y68 oryctolagus
1918	6	5.1	147	2	Q4HAR9_9DEIO	Q4har9 deinococcus	1991	6	5.1	154	2	Q9CAL7_ARATH	Q9cal7 arabidopsis
1919	6	5.1	147	2	Q6KEX4_METTU	Q6kex4 methylocell	1992	6	5.1	154	2	Q7PMN7_SYTH	Q7pmn7 symbiobact
1920	6	5.1	147	2	Q6KEX3_METPL	Q6kex3 methylocell	1993	6	5.1	154	2	P73658_SYNY3	P73658 synechocyst
1921	6	5.1	147	2	Q5P196_AZOSE	Q5p196 azoarcus ap	1994	6	5.1	154	2	Q7NL10_GLOVI	Q7nl10 gloeobacter
1922	6	5.1	147	2	Q7NRH0_CHRVO	Q7nrh0 chromobacte	1995	6	5.1	154	2	Q6K077_9PASS	Q6k077 cosaypha an
1923	6	5.1	147	2	Q9MJ37_9NEOB	Q9mj37 rana porosa	1996	6	5.1	154	2	Q98VX9_9HIV1	Q98vx9 human immun
1924	6	5.1	147	2	Q9MJ40_9NEOB	Q9mj40 rana porosa	1997	6	5.1	155	2	Q4J6Q5_SULAC	Q4j6q5 sulfolobus
1925	6	5.1	147	2	Q9MJ38_9NEOB	Q9mj38 rana porosa	1998	6	5.1	155	2	Q9HHI7_HALSA	Q9hh17 halobacteri
1926	6	5.1	147	2	Q9MJ41_RANNI	Q9mj41 rana nigrom	1999	6	5.1	155	2	Q9JPF3_NEIME	Q9jpf3 neisseria m
1927	6	5.1	147	2	Q9MJ39_9NEOB	Q9mj39 rana porosa	2000	6	5.1	155	2	Q5NUT6_9ACTO	Q5nut6 streptomyce
1928	6	5.1	148	2	Q6BSD1_DEBHA	Q6bsd1 debaryomyce							
1929	6	5.1	148	2	Q5P054_AZOSE	Q5p054 azoarcus ap							

ALIGNMENTS

```
RESULT 1
Q81594_9HEPC
ID Q81594_9HEPC PRELIMINARY; PRT; 209 AA.
AC Q81594_9HEPC PRELIMINARY; PRT; 209 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein 4 (Fragment).
GN Name=NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93249436; PubMed=7683463;
RA Stuyver L., Arnhem W.V., Wyeur A., Deleys R., Maertens G.;
RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV
type 3.";
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).
DR EMBL; D14600; BAA03449.1; -; Genomic_RNA.
DR PIR; PC1306; PCL306.
DR HSSP; P26663; 1CUI.
DR SMR; Q81594; 1-102.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 209
SQ SEQUENCE 209 AA; 23408 MW; 76648D9BB1D3CD12 CRC64;

Query Match 70.3%; Score 83; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 1.4e-72;
Matches 83; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGKPAIVPDKEVLYQQYD 60
Db 92 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGKPAIVPDKEVLYQQYD 151

QY 61 EMECSQAAPYIEQAQVIAHQPK 83
Db 152 EMECSQAAPYIEQAQVIAHQPK 174

RESULT 2
Q81595_9HEPC
ID Q81595_9HEPC PRELIMINARY; PRT; 133 AA.
AC Q81595_9HEPC PRELIMINARY; PRT; 133 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein 4 (Fragment).
GN Name=NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93249436; PubMed=7683463;
RA Stuyver L., Arnhem W.V., Wyeur A., Deleys R., Maertens G.;
RT "Analysis of the putative E1 envelope and NS4a epitope region of HCV
type 3.";
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).
DR EMBL; D14602; BAA03451.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
```

```
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 133
SQ SEQUENCE 133 AA; 14781 MW; 4BFF2128FD301691 CRC64;

Query Match 39.8%; Score 47; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 1.6e-37;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGKPA 47
Db 16 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVWIVGHIELGKPA 62

RESULT 3
O92933_9HEPC
ID O92933_9HEPC PRELIMINARY; PRT; 3021 AA.
AC O92933_9HEPC PRELIMINARY; PRT; 3021 AA.
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN-type 3a;
RA Shukla D.D., Chaturvedi S., Cao J.Y., Hoynes P.A.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF046866; AAC03058.1; -; Genomic_RNA.
DR HSSP; O8JY81; 1CW.
DR SMR; O92933; 1035-1663, 2431-2996.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:000236; F:serine-type peptidase activity; IEA.
DR GO; GO:000508; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:vital genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD/DEAH_N.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR Polyprotein.
SQ SEQUENCE 3021 AA; 328905 MW; D7B6133B33030303CD CRC64;
```

```
Query Match      39.8%; Score 47; DB 2; Length 3021;
Best Local Similarity 100.0%; Pred. No. 1.9e-36;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPA 47
      |||||
DB      1653 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPA 1699

RESULT 4
Q68870_9HEPC PRELIMINARY; PRT; 3021 AA.
AC Q68870;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Genes for core, envelope and NS1 proteins.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Tissue-Serum;
RA Seelig R., Weber P., Seeling H.P., Ledger N., Bottner C., Renz M.;
RT "Hepatitis C virus type V genome isolated from a patient in Germany.";
RL Submitted (JAN-1995) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93249436; PubMed=7683463;
RA Stuyver L., Arnhem W.V., Wyseur A., Deleys R., Maertens G.;
RT "Analysis of the putative E1 envelope A, and NS4a epitope region of HCV
RT type 3.";
RL Biochem. Biophys. Res. Commun. 192:635-641(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92268871; PubMed=1316939;
RA Chan S., McOmish P., Holmes E., Dow B., Peutherer J., Follett E.,
RA Yap P., Simmonds P.;
RT "Analysis of a new hepatitis C virus type and its phylogenetic
RT relationship to existing variants.";
RL J. Gen. Virol. 73:1131-1141(1992).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dubeiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; X76918; CAA54244.1; -; mRNA.
DR PIR; PC1307; PC1307.
DR PIR; PQ0401; PQ0401.
DR PIR; PQ0804; PQ0804.
DR PIR; S41288; S41288.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q68870; 1035-1663, 2431-2996.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; F:viral genome replication; IEA.
DR GO; GO:0019087; F:viral genome reformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
```

```
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_P8.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4; 1.
DR Pfam; PF01001; HCV_NS4a; 1.
DR Pfam; PF01506; HCV_NS4b; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXdc; 1.
KW Envelope protein.
FT CHAIN 1 191
FT CHAIN 192 383
FT CHAIN 384 735
FT CHAIN 736 1012
FT CHAIN 1013 1663
FT CHAIN 1664 1717
FT CHAIN 1718 1978
FT CHAIN 1979 2430
FT CHAIN 2431 3021
SQ SEQUENCE 3021 AA; 329096 MW; BF2B499AA5A58CF CRC64;

Query Match      39.8%; Score 47; DB 2; Length 3021;
Best Local Similarity 100.0%; Pred. No. 1.9e-36;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPA 47
      |||||
DB      1653 ACMSADLEVTSTWVLLGGVLAALAAAYCLSVGCVVIVGHIELGGKPA 1699

RESULT 5
Q81258_9HEPC PRELIMINARY; PRT; 3021 AA.
AC Q81258;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NZL1;
RA Sakamoto M.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NZL1;
RA Sakamoto H.;
RL Submitted (SEP-1993) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92268871; PubMed=1316939;
RA Chan S., McOmish P., Holmes E., Dow B., Peutherer J., Follett E.,
RA Yap P., Simmonds P.;
RT "Analysis of a new hepatitis C virus type and its phylogenetic
RT relationship to existing variants.";
RN [4]
```


J. Gen. Virol. 73:1131-1141 (1992).

[4]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=9324886; PubMed=8385694;

RA Simmonds P., McMahon F., Yap P.L., Chan S.-W.W., Lin C.K.,

RA Dushenko G., Saeed A.A., Holmes E.C.;

RT "Sequence variability in the 5' non-coding region of hepatitis C

RT virus: identification of a new virus type and restrictions on sequence

RT diversity.";

RL J. Gen. Virol. 74:661-668 (1993).

DR EMBL; D17763; BAA04609.1; -; Genomic_RNA.

DR FIR; PQ0401; PQ0401.

DR PIR; PQ0804; PQ0804.

DR HSP; Q8JYS1; 1CWX.

DR SMR; Q81258; 1035-1663; 2431-2996.

DR GO; GO:0019028; C:Viral capsid; IEA.

DR GO; GO:0019031; C:Viral envelope; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0003723; F:RNA binding; IEA.

DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.

DR GO; GO:0008236; F:serine-type peptidase activity; IEA.

DR GO; GO:0005198; F:structural molecule activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR GO; GO:0006350; P:transcription; IEA.

DR GO; GO:0019079; P:Viral genome replication; IEA.

DR GO; GO:0019087; P:Viral transformation; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR011545; DEAD/DEAH N.

DR InterPro; IPR002522; HCV capsid.

DR InterPro; IPR002521; HCV core.

DR InterPro; IPR002519; HCV env.

DR InterPro; IPR002531; HCV NS1.

DR InterPro; IPR000745; HCV NS4a.

DR InterPro; IPR001490; HCV NS4b.

DR InterPro; IPR002868; HCV NS5a.

DR InterPro; IPR002166; HCV RdRp.

DR InterPro; IPR002518; Pept_U39 HCV NS2.

DR InterPro; IPR004109; Peptidase_S29.

DR InterPro; IPR002016; Peroxidase.

DR InterPro; IPR007095; RNA pol DS PS.

DR InterPro; IPR007094; RNA pol_PSVir.

DR Pfam; PF01543; HCV capsid; 1.

DR Pfam; PF01542; HCV core; 1.

DR Pfam; PF01539; HCV env; 1.

DR Pfam; PF01560; HCV NS1; 1.

DR Pfam; PF01538; HCV NS2; 1.

DR Pfam; PF02907; HCV NS3; 1.

DR Pfam; PF01006; HCV NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.

DR Pfam; PF01506; HCV NS5a; 1.

DR Pfam; PF00998; RdRp 3; 1.

DR SMART; SM00487; DEXDC; 1.

DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.

KW Polyprotein.

FT CHAIN 1 191 C.

FT CHAIN 192 383 E1.

FT CHAIN 384 735 E2/NS1.

FT CHAIN 736 1012 NS2.

FT CHAIN 1013 1663 NS3.

FT CHAIN 1664 1717 NS4a.

FT CHAIN 1718 1978 NS4b.

FT CHAIN 1979 2430 NS5a.

FT CHAIN 2431 3021 NS5b.

SQ SEQUENCE 3021 AA; 329578 MW; 38712CCBC0C19562 CRC64;

Query Match 35.6%; Score 42; DB 2; Length 3021;

Best Local Similarity 100.0%; Pred. No. 1.5e-31;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVIVGHIEL 42

DB 1653 ACMSADLEVTSTWLLGGVLAALAAAYCLSVGCVIVGHIEL 1694

RESULT 6

O39899_9HEPC PRELIMINARY; PRT; 70 AA.

AC O39899_9HEPC PRELIMINARY; PRT; 70 AA.

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Non-structural protein 4a/b (Fragment).

GN Name=NS4a/b;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OC NCBI_TaxID=11103;

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=98032593; PubMed=9365889;

RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;

RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,

RA Simmonds P.;

RT "Sequence analysis of hepatitis C virus variants producing discrepant

RT results with two different genotyping assays.";

RL J. Med. Virol. 53:237-244 (1997).

DR EMBL; AF007504; AAB62955.2; -; Genomic_RNA.

DR InterPro; IPR000745; HCV NS4a.

DR Pfam; PF01006; HCV NS4a; 1.

FT NON_TER 1 1

FT NON_TER 70 70

SQ SEQUENCE 70 AA; 7743 MW; 818D296E0E48DB1 CRC64;

Query Match 29.7%; Score 35; DB 2; Length 70;

Best Local Similarity 100.0%; Pred. No. 5.1e-26;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQYDMEECSSQAAPYIEQAQVIAHQFK 83

DB 22 VPDKEVLYQYDMEECSSQAAPYIEQAQVIAHQFK 56

RESULT 7

O39898_9HEPC PRELIMINARY; PRT; 88 AA.

AC O39898_9HEPC PRELIMINARY; PRT; 88 AA.

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Non-structural protein 4a/b (Fragment).

GN Name=NS4a/b;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OC NCBI_TaxID=11103;

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=98032593; PubMed=9365889;

RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;

RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,

RA Simmonds P.;

RT "Sequence analysis of hepatitis C virus variants producing discrepant

RT results with two different genotyping assays.";

RL J. Med. Virol. 53:237-244 (1997).

DR EMBL; AF007503; AAB62954.2; -; Genomic_RNA.

DR InterPro; IPR000745; HCV NS4a.

DR InterPro; IPR001490; HCV NS4b.

DR Pfam; PF01006; HCV NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.

FT NON_TER 1 1

FT NON_TER 88 88

SQ SEQUENCE 88 AA; 9838 MW; 91FDFBC4EDD171FF CRC64;

Query Match 29.7%; Score 35; DB 2; Length 88;

Best Local Similarity 100.0%; Pred. No. 6.1e-26;

```
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 83
    |||||
Db 23 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 57

RESULT 8
C39904_9HEPC
ID C39904_9HEPC PRELIMINARY; PRT; 90 AA.
AC C39904;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007509; AAB62962.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10000 MW; 15465423A211108B CRC64;

Query Match 29.7%; Score 35; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 83
    |||||
Db 25 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 59

RESULT 9
C39906_9HEPC
ID C39906_9HEPC PRELIMINARY; PRT; 90 AA.
AC C39906;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007511; AAB62962.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
```

```
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 10031 MW; DC7CDBE68881109C CRC64;

Query Match 29.7%; Score 35; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 83
    |||||
Db 25 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 59

RESULT 10
C39908_9HEPC
ID C39908_9HEPC PRELIMINARY; PRT; 90 AA.
AC C39908;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007513; AAB62964.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 9982 MW; D4E5CD86887769C CRC64;

Query Match 29.7%; Score 35; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 83
    |||||
Db 25 VPDKEVLVQYDMEECQAAPYIEQAQVIAHQFK 59

RESULT 11
C39910_9HEPC
ID C39910_9HEPC PRELIMINARY; PRT; 90 AA.
AC C39910;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
```

```
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RL results with two different genotyping assays.";
DR J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007515; AAB62966.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR NON_TER 1
FT NON_TER 90
SQ SEQUENCE 90 AA; 9982 MW; D4E5CD868E87769C CRC64;

Query Match 29.7%; Score 35; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 6.2e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLVQQYDEMECSQAAPYIEQAQVIAHQFK 83
DB 25 VPDKEVLVQQYDEMECSQAAPYIEQAQVIAHQFK 59

RESULT 12
Q68239_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68239;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ja;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yip K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14275; AAC53964.1; -; Genomic_RNA.
DR HSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15311 MW; B45AC0P8917DBAAC CRC64;

Query Match 29.7%; Score 35; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 8.8e-26;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLVQQYDEMECSQAAPYIEQAQVIAHQFK 83
DB 69 VPDKEVLVQQYDEMECSQAAPYIEQAQVIAHQFK 103

RESULT 13
Q39900_9HEPC PRELIMINARY; PRT; 83 AA.
AC Q39900;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
```

```
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007505; AAB62956.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR NON_TER 1
FT NON_TER 83
SQ SEQUENCE 83 AA; 9089 MW; 0FAA36D622B24BDE CRC64;

Query Match 27.1%; Score 32; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 5e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 KEVLYQQYDEMECSQAAPYIEQAQVIAHQFK 83
DB 28 KEVLYQQYDEMECSQAAPYIEQAQVIAHQFK 59

RESULT 14
Q39896_9HEPC PRELIMINARY; PRT; 84 AA.
AC Q39896;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007501; AAB62952.2; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR NON_TER 1
FT NON_TER 84
SQ SEQUENCE 84 AA; 9343 MW; 135FFAB56ADD4D16 CRC64;

Query Match 26.3%; Score 31; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.8e-22;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQYDEMECSQAAPYIEQAQVIAHQFK 83
DB 28 EVLYQQYDEMECSQAAPYIEQAQVIAHQFK 58

RESULT 15
Q39917_9HEPC PRELIMINARY; PRT; 84 AA.
AC Q39917;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
```

```

DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP NUCLEOTIDE SEQUENCE.
RP
RX MDLINE=98032593; PubMed=9365989;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007522; AAB62973.2; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 84
FT STOP 84
SQ SEQUENCE 84 AA; 9360 MW; 88E9F6DBE5E66DA7 CRC64;

Query Match 26.3%; Score 31; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.9e-22;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 EVLYQQYDEMECSQAAPYIEQAQVIAHQFK 83
DB 24 EVLYQQYDEMECSQAAPYIEQAQVIAHQFK 54
|||||
|||||

RESULT 16
Q81487_9HEPC
ID Q81487_9HEPC PRELIMINARY; PRT; 3023 AA.
AC Q81487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus type 3b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=42791;
[1]
RP NUCLEOTIDE SEQUENCE.
RP
RA Chayana K.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; D49374; BAA08372.1; -; Genomic_RNA.
DR HSP; P27958; 1A1V.
DR SMR; Q81487; 1037-1665, 2433-2998.
DR MEQPS; S29.001; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.

```

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 77
Db 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 53

RESULT 18
O39894_9HEPC
ID O39894_9HEPC PRELIMINARY; PRT; 89 AA.
AC O39894;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007499; AAB62950.1; -; Genomic_RNA.
DR InterPro; IPR001490; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 1
FT NON_TER 89 89
SQ SEQUENCE 89 AA; 9805 MW; 31D53EA815CF6F1P CRC64;

Query Match 24.6%; Score 29; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 77
Db 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 53

RESULT 19
O39905_9HEPC
ID O39905_9HEPC PRELIMINARY; PRT; 90 AA.
AC O39905;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007510; AAB62961.1; -; Genomic_RNA.
DR InterPro; IPR001490; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 1
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 9888 MW; DC63CBE6FB32B0C CRC64;

Query Match 24.6%; Score 29; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FT NON_TER 90 90
SQ SEQUENCE 90 AA; 9988 MW; DC63CBE6FB32B0C CRC64;

Query Match 24.6%; Score 29; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.6e-20;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 77
Db 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 53

RESULT 20
O39903_9HEPC
ID O39903_9HEPC PRELIMINARY; PRT; 84 AA.
AC O39903;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007508; AAB62959.1; -; Genomic_RNA.
DR InterPro; IPR001490; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 1
FT NON_TER 84 84
SQ SEQUENCE 84 AA; 9228 MW; E2EF3688337C5D45 CRC64;

Query Match 23.7%; Score 28; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 4.1e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 76
Db 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQV 52

RESULT 21
O39901_9HEPC
ID O39901_9HEPC PRELIMINARY; PRT; 90 AA.
AC O39901;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";

RL J. Med. Virol. 53:237-244 (1997).
DR EMBL; AF007506; AAB62957.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 1 90
SQ SEQUENCE 90 AA; 10055 MW; B6D13103900A75C7 CRC64;

Query Match 23.7%; Score 28; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 76
DB 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 52

RESULT 22

ID O39902_9HEPC PRELIMINARY; PRT; 90 AA.
AC O39902;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RA DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10<3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244 (1997).
DR EMBL; AF007507; AAB62958.1; -; Genomic_RNA.
DR InterPro; IPR007507; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 90
SQ SEQUENCE 90 AA; 10016 MW; A5F1DF3D69BA5285 CRC64;

Query Match 23.7%; Score 28; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 76
DB 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 52

RESULT 23

ID O39912_9HEPC PRELIMINARY; PRT; 90 AA.
AC O39912;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;

RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10<3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244 (1997).
DR EMBL; AF007517; AAB62968.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 90
SQ SEQUENCE 90 AA; 9988 MW; D8FD53A698577E59 CRC64;

Query Match 23.7%; Score 28; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 76
DB 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 52

RESULT 24

ID O39913_9HEPC PRELIMINARY; PRT; 90 AA.
AC O39913;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RA DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10<3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244 (1997).
DR EMBL; AF007518; AAB62969.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 90
SQ SEQUENCE 90 AA; 10000 MW; C7CDF95B15A6A40 CRC64;

Query Match 23.7%; Score 28; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 76
DB 25 VPDKEVLYQQYDEMEECSSQAAPYIEQAQ 52

RESULT 25

ID O39916_9HEPC PRELIMINARY; PRT; 90 AA.
AC O39916;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name=NS4a/b;

```
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007521; AB62972.1; -; Genomic_RNA.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 90
FT NON_TER 90
SQ SEQUENCE 90 AA; 9972 MW; C7DF7495B15A6A40 CRC64;

Query Match 23.7%; Score 28; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 4.4e-19;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 VPDKEVLVQYDEMECSQAAPYIEQAQ 76
DB 25 VPDKEVLVQYDEMECSQAAPYIEQAQ 52

RESULT 26
Q68233_9HEPC
ID Q68233_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68233;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3a;
RX MEDLINE=95146953; PubMed=7844535;
RX Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RA InterPro; IPR000745; HCV_NS4a.
RA Pfam; PF01006; HCV_NS4a; 1.
RA Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
FT NON_TER 138
SQ SEQUENCE 138 AA; 15379 MW; 45236C0E5427B19F CRC64;

Query Match 19.5%; Score 23; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWVLGGVLA 23
DB 21 ACMSADLEVTSTWVLGGVLA 43

RESULT 27
Q68241_9HEPC
ID Q68241_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68241;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=3a;
RX MEDLINE=95146953; PubMed=7844535;
RX Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RA InterPro; IPR000745; HCV_NS4a.
RA Pfam; PF01006; HCV_NS4a; 1.
RA Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
FT NON_TER 138
SQ SEQUENCE 138 AA; 15277 MW; 275F2F928F5A10E4 CRC64;

Query Match 19.5%; Score 23; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWVLGGVLA 23
DB 21 ACMSADLEVTSTWVLGGVLA 43

RESULT 28
Q81495_9HEPC
ID Q81495_9HEPC PRELIMINARY; PRT; 3021 AA.
AC Q81495;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K3a;
RX MEDLINE=95053917; PubMed=7964640;
RX Yamada N., Manihara K., Mizokami M., Ohba K., Takada A., Tsutsumi M.,
RA Date T.;
RT "Full-length sequence of the genome of hepatitis C virus type 3a:
RT comparative study with different genotypes.";
RL J. Gen. Virol. 75:3279-3284(1994).
DR EMBL; D28917; BAA06044.1; -; Genomic_RNA.
DR HSSP; P26664; 1HEI.
DR SMR; Q81495; 1035-1663, 2431-2996.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
```

DR GO: 0019087; P: Viral transformation; IEA.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR011545; DEAD/DEAH N.

DR InterPro; IPR002522; HCV capsid.

DR InterPro; IPR002521; HCV core.

DR InterPro; IPR002519; HCV env.

DR InterPro; IPR002531; HCV NS1.

DR InterPro; IPR000745; HCV NS4a.

DR InterPro; IPR001490; HCV NS4b.

DR InterPro; IPR002868; HCV NS5a.

DR InterPro; IPR002166; HCV NS5a.

DR InterPro; IPR004109; Peptidase S29.

DR InterPro; IPR002518; Pept_U39 HCV NS2.

DR InterPro; IPR007095; RNA_pol_PS.

DR InterPro; IPR007094; RNA_pol_Psvir.

DR Pfam; PF01543; HCV capsid; 1.

DR Pfam; PF01542; HCV core; 1.

DR Pfam; PF01539; HCV env; 1.

DR Pfam; PF01560; HCV NS1; 1.

DR Pfam; PF01538; HCV NS2; 1.

DR Pfam; PF02907; HCV NS3; 1.

DR Pfam; PF01006; HCV NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.

DR Pfam; PF01506; HCV NS5a; 1.

DR Pfam; PF00998; RdRP 3; 1.

DR SMART; SM00487; DEXDC; 1.

KW Polyprotein.

SQ SEQUENCE 3021 AA; 328387 MW; A97418FF36C062A4 CRC64;

Query Match 19.5%; Score 23; DB 2; Length 3021;

Best Local Similarity 100.0%; Pred. No. 5.6e-13;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACMSADLEVTSTWVLGGVLA 23

Db 1653 ACMSADLEVTSTWVLGGVLA 1675

RESULT 29

O39909_9HEPC

ID O39909_9HEPC PRELIMINARY; PRT; 90 AA.

AC O39909

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Non-structural protein 4a/b (Fragment).

GN Name=NS4a/b;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=111103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=98032593; PubMed=9365889;

RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;

RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,

RA Simmonds P.;

RT "Sequence analysis of hepatitis C virus variants producing discrepant

results with two different genotyping assays.";

RL J. Med. Virol. 53:237-244(1997).

DR EMBL; AF007514; AB62965.1; -; Genomic_RNA.

DR InterPro; IPR000745; HCV_NS4a.

DR InterPro; IPR001490; HCV_NS4b.

DR Pfam; PF01006; HCV NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.

FT NON_TER 1

FT NON_TER 90

SQ SEQUENCE 90 AA; 10013 MW; D4E7478681862B30 CRC64;

Query Match

Best Local Similarity 18.6%; Score 22; DB 2; Length 90;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 AYCLSVGCVVIVGHIELGGKPA 47

Db 2 AYCLSVGCVVIVGHIELGGKPA 23

RESULT 30

O39907_9HEPC

ID O39907_9HEPC PRELIMINARY; PRT; 87 AA.

AC O39907

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Non-structural protein 4a/b (Fragment).

GN Name=NS4a/b;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=111103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=98032593; PubMed=9365889;

RX DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;

RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,

RA Simmonds P.;

RT "Sequence analysis of hepatitis C virus variants producing discrepant

results with two different genotyping assays.";

RL J. Med. Virol. 53:237-244(1997).

DR EMBL; AF007512; AB62963.1; -; Genomic_RNA.

DR InterPro; IPR000745; HCV_NS4a.

DR InterPro; IPR001490; HCV_NS4b.

DR Pfam; PF01006; HCV NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.

FT NON_TER 1

FT NON_TER 87

SQ SEQUENCE 87 AA; 9692 MW; EB295E968DD9B8F8 CRC64;

Query Match

Best Local Similarity 17.8%; Score 21; DB 2; Length 87;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 27 YCLSVGCVVIVGHIELGGKPA 47

Db 3 YCLSVGCVVIVGHIELGGKPA 23

RESULT 31

O56637_9HEPC

ID O56637_9HEPC PRELIMINARY; PRT; 193 AA.

AC O56637

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Non-structural protein (Fragment).

GN Name=NS4;

OS Hepatitis C virus type 3g.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=42792;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN-type 3g;

RA Panigrahi A.K., Panda S.K.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF042791; AB97852.1; -; Genomic_RNA.

DR HSSP; P26663; 1CU1.

DR SMR; O56637; 1-85.

DR MEROPS; S29.001; -.

DR InterPro; IPR000745; HCV_NS4a.

DR InterPro; IPR001490; HCV_NS4b.

DR Pfam; PF01006; HCV NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.

FT NON_TER 1

FT NON_TER 193


```
SQ SEQUENCE 193 AA; 21493 MW; AF7B8E3FB3B69505 CRC64;
Query Match 16.9%; Score 20; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLAALAAAYCLSVGCVWI 36
Db 91 LGGVLAALAAAYCLSVGCVWI 110

RESULT 32
Q4QTD9_9HEPC
ID Q4QTD9_9HEPC PRELIMINARY; PRT; 3015 AA.
AC Q4QTD9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KM45;
RX PubMed=15847930; DOI=10.1016/j.jviromet.2005.01.031;
RA Lu L., Nakano T., Smallwood G.A., Heffron T.G., Robertson B.H.,
RA Hagedorn C.H.;
RT "A refined long RT-PCR technique to amplify complete viral RNA genome
RT sequences from clinical samples: application to a novel hepatitis C
RT virus variant of genotype 6."
RL J. Virol. Methods 126:139-148(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KM45;
RA Lu L., Nakano T., Fu Y., Miller S., Kuiken C., Negro F.,
RA Hagedorn C.H., Robertson B.H.;
RT "Genome sequences of four hepatitis C virus variants from China
RT representing new genotype 6 members and a novel subtype."
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY878650; AAX81527.1; -; Genomic_RNA.
KW Polyprotein.
SQ SEQUENCE 3015 AA; 327970 MW; 84FES18175A0396B CRC64;

Query Match 16.1%; Score 19; DB 2; Length 3015;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLSVGCVWI 36
Db 1665 GGVLAALAAAYCLSVGCVWI 1693

RESULT 33
O39927_9HEPC
ID O39927_9HEPC PRELIMINARY; PRT; 3018 AA.
AC O39927;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus type 6a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31655;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=euhk2;
RX MEDLINE=97320431; PubMed=9177282; DOI=10.1006/bbrc.1997.6627;
RA Adams A., Chamberlain R.W., Taylor L.A., Davidson F., Lin C.K.,
RA Simmonds P., Elliot R.M.;
RT "Complete coding sequence of hepatitis C virus genotype 6a.";
```

```
RL Biochem. Biophys. Res. Commun. 234:393-396(1997).
DR EMBL; Y12083; CAA72801.1; -; Genomic_RNA.
DR HSP; O8JYS1; 1CWX.
DR SNR; O39927; 1033-1661, 2428-2993.
DR MEROPS; C18.001; -.
DR MEROPS; S29.001; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV-env.
DR InterPro; IPR002531; HCV-NS1.
DR InterPro; IPR000745; HCV-NS4a.
DR InterPro; IPR001490; HCV-NS4b.
DR InterPro; IPR002868; HCV-NS5a.
DR InterPro; IPR002166; HCV-RdRp.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV-env; 1.
DR Pfam; PF01560; HCV-NS1; 1.
DR Pfam; PF01538; HCV-NS2; 1.
DR Pfam; PF02907; HCV-NS3; 1.
DR Pfam; PF01006; HCV-NS4a; 1.
DR Pfam; PF01001; HCV-NS4b; 1.
DR Pfam; PF01506; HCV-NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3018 AA; 329017 MW; 6B67FB3CF6A61AE3 CRC64;

Query Match 16.1%; Score 19; DB 2; Length 3018;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLSVGCVWI 36
Db 1668 GGVLAALAAAYCLSVGCVWI 1686

RESULT 34
O512N3_9HEPC
ID O512N3_9HEPC PRELIMINARY; PRT; 3019 AA.
AC O512N3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus type 6a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31655;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=6a33;
RA Zhou X.W., Chan P.K.S., Tam J.S.L.;
RT "Full genome sequence of HCV 6a virus strains isolated in Hong Kong."
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
```

DR EMBL, AY859526; AAM56714.1; -, Genomic_RNA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol_D8_P8.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
FT CHAIN 1 191 putative core protein.
FT CHAIN 192 383 putative E1 protein.
FT CHAIN 384 751 putative short E2 protein.
FT CHAIN 752 814 putative p7 protein.
FT CHAIN 815 1031 putative NS2 protein.
FT CHAIN 1032 1662 putative NS3 protein.
FT CHAIN 1663 1716 putative NS4 protein.
FT CHAIN 1717 1977 putative NS4B protein.
FT CHAIN 1978 2428 putative NS5A protein.
FT CHAIN 2429 3019 putative NS5B protein.
SQ SEQUENCE 3019 AA; 328853 MW; FFL161164B164DF3 CRC64;

Query Match 16.1%; Score 19; DB 2; Length 3019;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCISVGCVV 36
|||||
DB 1669 GGVLAALAAAYCISVGCVV 1687

RESULT 35
O92531.9HEPC
ID O92531.9HEPC PRELIMINARY; PRT; 3016 AA.
AC O92531;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VN405;
RX MEDLINE=98378034; PubMed=9714232;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tauda F., Miyakawa Y.,
RA Mayumi M.;
RT "The entire nucleotide sequences of three hepatitis C virus isolates
in genetic groups 7-9 and comparison with those in the other eight
genetic groups.";
RL J. Gen. Virol. 79:1847-1857(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
virus: identification of a new virus type and restrictions on sequence
diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; D84264; BAA32666.1; -, Genomic_RNA.
DR FIR; PQ0804; PQ0804.
DR HSBP; Q8JYS1; ICWX.
DR SMR; O92531; 1031-1659, 2426-2955.
DR MEROPS; C18.001; -.
DR MEROPS; S29.001; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol_D8_P8.
DR InterPro; IPR007094; RNA pol_P5vir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3016 AA; 328035 MW; 4B5CFF96258BCE3B CRC64;

Query Match 15.3%; Score 18; DB 2; Length 3016;
Best Local Similarity 100.0%; Pred. No. 4.3e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCISVGCVV 35
|||||
DB 1666 GGVLAALAAAYCISVGCVV 1683

```
RESULT 36
Q68801_9HEPC
ID Q68801_9HEPC PRELIMINARY; PRT; 3019 AA.
AC Q68801;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96226020; PubMed=8627233;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F.,
RA Lesmana L.A., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Jakarta, Indonesia classifiable into
RT novel genotypes in the second (2e and 2f), tenth (10a) and eleventh
RT (11a) genetic groups.";
RL J. Gen. Virol. 77:293-301(1996).
DR EMBL; D63821; BAA09890.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q68801; 1034-1662, 2429-2994.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001917; AminoTrans II.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF01506; HCV NS4a; 1.
DR Pfam; PF01507; HCV NS4b; 1.
DR Pfam; PF01508; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00599; AA_TRANSPR_CLASS_2; UNKNOWN_1.
KW Polyprotein.
SQ
SEQUENCE 3019 AA; 32821 MW; AF7A6774BC6D95FA CRC64;
Query Match 15.3%; Score 18; DB 2; Length 3019;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 49 VPDKEVLYQQYDEMEBCS 66
DB 1700 VPDKEVLYQQYDEMEBCS 1717
```

```
RESULT 37
O39929_9HEPC
ID O39929_9HEPC PRELIMINARY; PRT; 3008 AA.
AC O39929;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE HCV polyprotein.
OS Hepatitis C virus type 4a.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31653;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97335261; PubMed=9191927;
RA Chamberlain R.W., Adams N., Saeed A.A., Simmonds P., Elliott R.M.;
RT "Complete nucleotide sequence of a type 4 hepatitis C virus variant,
RT the predominant genotype in the Middle East.";
RL J. Gen. Virol. 78:1341-1347(1997).
DR EMBL; Y11604; CA72338.1; -; Genomic_RNA.
DR FIR; PQ0804; PQ0804.
DR HSP; Q8JYS1; 1CWX.
DR SMR; O39929; 1029-1657, 2418-2982.
DR MEROPS; S29.001; -.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF01506; HCV NS4a; 1.
DR Pfam; PF01507; HCV NS4b; 1.
DR Pfam; PF01508; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00599; AA_TRANSPR_CLASS_2; UNKNOWN_1.
KW Polyprotein.
SQ
SEQUENCE 3008 AA; 32821 MW; AF7A6774BC6D95FA CRC64;
Query Match 15.3%; Score 18; DB 2; Length 3008;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 49 VPDKEVLYQQYDEMEBCS 66
DB 1700 VPDKEVLYQQYDEMEBCS 1717
```

```
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Polyprotein.
SQ SEQUENCE 3008 AA; 327599 MW; 8E7FC932E27C406F CRC64;

Query Match 12.7%; Score 15; DB 2; Length 3008;
Best Local Similarity 100.0%; Pred.No. 3.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLSVG 32
Db 1664 GGVLAALAAAYCLSVG 1678

RESULT 38
Q92532_9HEPC
ID O92532_9HEPC PRELIMINARY; PRT; 3015 AA.
AC O92532;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VN004;
RX MEDLINE=98378034; PubMed=9714232;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F., Miyakawa Y.,
RA Mayumi M.;
RT "The entire nucleotide sequences of three hepatitis C virus isolates
RT in genetic groups 7-9 and comparison with those in the other eight
RT genetic groups.";
RL J. Gen. Virol. 79:1847-1857(1998).
DR EMBL; D84265; BAA32667.1; -; Genomic_RNA.
DR HSP; Q8JVS1; ICWK.
DR SMR; Q92532; 1030-1658, 2425-2954.
DR MEROPS; C18.001; -.
DR MEROPS; S29.001; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; F:viral genome replication; IEA.
DR GO; GO:0019087; F:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.

DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Polyprotein.
SQ SEQUENCE 3015 AA; 327975 MW; 69108DD32B5DA012 CRC64;

Query Match 12.7%; Score 15; DB 2; Length 3015;
Best Local Similarity 100.0%; Pred.No. 3.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYCLSVGCVVI 36
Db 1669 AALAAAYCLSVGCVVI 1683

RESULT 39
Q68222_9HEPC
ID Q68222_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68222;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=951146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14258; AAC53947.1; -; Genomic_RNA.
DR HSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15298 MW; 6389770EB60F67B6 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred.No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
Db 38 GGVLAALAAAYCLIS 50

RESULT 40
Q68223_9HEPC
ID Q68223_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68223;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (fragment).
GN Name=ns4;
```

```
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14259; AAC53948.1; -; Genomic_RNA.
DR HSSP; P27958; 1HEI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15392 MW; F8CB866A53AA907B CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
38 GGVLAALAAAYCLS 50

RESULT 41
Q68224_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68224;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=n84;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14260; AAC53949.1; -; Genomic_RNA.
DR HSSP; P27958; 1HEI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15295 MW; 01335FB49A841A53 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
38 GGVLAALAAAYCLS 50

RESULT 42
Q68225_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68225;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=n84;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14261; AAC53950.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15364 MW; C23FDA160A9B0A41 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
38 GGVLAALAAAYCLS 50

RESULT 43
Q68226_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68226;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=n84;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14262; AAC53951.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15309 MW; C375FDEB9B376A41 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
ID Q68225_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68225;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=n84;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14261; AAC53950.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15364 MW; C23FDA160A9B0A41 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB |||||
38 GGVLAALAAAYCLS 50

RESULT 43
Q68226_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68226;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=n84;
OS Hepatitis C virus.
OC Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
  Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14262; AAC53951.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15309 MW; C375FDEB9B376A41 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 18 GGVLAALAAAYCCLS 30
Db 38 GGVLAALAAAYCCLS 50

RESULT 44
Q68236_9HEPC
ID Q68236_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68236;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14272; AAC53961.1; -; Genomic_RNA.
DR HSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 138 AA; 15342 MW; 5089ECBB9A840906 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCCLS 30
Db 38 GGVLAALAAAYCCLS 50

RESULT 45
Q68243_9HEPC
ID Q68243_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68243;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14279; AAC53968.1; -; Genomic_RNA.
DR HSP; P27958; 1HEI.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 138 AA; 15309 MW; E330195555B076DB CRC64;

Query Match 11.0%; Score 13; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCCLS 30
Db 38 GGVLAALAAAYCCLS 50

RESULT 46
Q02258_9HEPC
ID Q02258_9HEPC PRELIMINARY; PRT; 313 AA.
AC Q02258;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE ORF 1 (Fragment).
GN Name=NS3/NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92013977; PubMed=1655961;
RA Krensdorf D., Porchon C., Kim J.P., Reyes G.R., Brechot C.;
RT "Partial nucleotide sequence analysis of a French hepatitis C virus:
RT implications for HCV genetic variability in the E2/NS1 protein.";
RL J. Gen. Virol. 72:2557-2561 (1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94174722; PubMed=7510436;
RA Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
RA Inchauste G.;
RT "Characterization and mapping of a B-cell immunogenic domain in
RT hepatitis C virus E2 glycoprotein using a yeast peptide library.";
RL Virology 200:246-255 (1994).
DR EMBL; D10664; BAA01515.1; -; Genomic_RNA.
DR HSP; P27958; 1HEI.
DR SMR; Q02258; 1-201.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 313 AA; 35082 MW; BCE66DDA60F2D15B CRC64;

Query Match 11.0%; Score 13; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCCLS 30
Db 210 GGVLAALAAAYCCLS 222

RESULT 47
Q68K64_9HEPC
ID Q68K64_9HEPC PRELIMINARY; PRT; 636 AA.
AC Q68K64;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

```

```
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685571; AAT94249.1; -; Genomic_RNA.
DR SMR; Q68K54; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 636 636
FT SEQUENCE 636 AA; 68085 MW; CF80D9639795BACB CRC64;

Query Match 11.0%; Score 13; DB 2; Length 636;
Best Local Similarity 100.0%; Pred. No. 0.00097;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 48
Q68K51_9HEPC PRELIMINARY; PRT; 652 AA.
ID Q68K51;
AC Q68K51;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685564; AAT94247.1; -; Genomic_RNA.
DR SMR; Q68K66; 3-551.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 636 636
FT SEQUENCE 658 AA; 70676 MW; B1C0F9F7BFBAE4E3 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 49
Q68K66_9HEPC PRELIMINARY; PRT; 658 AA.
ID Q68K66;
AC Q68K66;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685564; AAT94247.1; -; Genomic_RNA.
DR SMR; Q68K66; 3-551.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 658 658
FT SEQUENCE 658 AA; 70676 MW; B1C0F9F7BFBAE4E3 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571
```


DR	SMART; SM00490; HELICc; 1.
KW	ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW	Structural protein.
FT	NON TER 1
FT	NON TER 658
SQ	SEQUENCE 658 AA; 70788 MW; 376935CDE6E193C90 CRC64;
Query Match	
Best Local Similarity 100.0%; Score 13; DB 2; Length 659;	
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	18 GGVLAAALAAAYCLS 30
Dd	558 GGVLAAALAAAYCLS 570
RESULT 53	
Q68K68_9HEPC	
ID	Q68K68_9HEPC PRELIMINARY; PRT; 659 AA.
AC	Q68K68;
DT	25-OCT-2004 (TrEMBLrel. 28, Created)
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DS	Polyprotein (Fragment).
OS	Hepatitis C virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC	Hepadnavirus.
OX	NCBI_TaxID=11103;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA	Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT	"Sequence Analysis of Hepatitis C Virus Replication Functions in
RT	RCV/HIV Coinfected Subjects";
RL	Submitted (JUL-2004) to the EMBL/GenBank/DBDJ databases.
DR	EMBL; AY585562; AA794245.1; -; Genomic_RNA.
DR	SMR; Q68K68; 1-552.
GO	GO:0005524; F:ATP binding; IEA.
DR	GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR	GO: GO:0003676; F:nucleic acid binding; IEA.
DR	GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR	GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR	GO: GO:0019087; P:viral transformation; IEA.
DR	InterPro; IPR001410; DEAD.
DR	InterPro; IPR011545; DEAD/DEAH_N.
DR	InterPro; IPR000745; HCV_NS4a_
DR	InterPro; IPR001490; HCV_NS4b.
DR	InterPro; IPR004109; Peptide_S29.
DR	Pfam; PF02907; HCV_NS3; 1.
DR	Pfam; PF01006; HCV_NS4a; 1.
DR	Pfam; PF01001; HCV_NS4b; 1.
DR	SMART; SM00487; DEXdC; 1.
KW	Capsid protein; Polyprotein; Structural protein.
FT	NON TER 1
FT	NON TER 659
SQ	SEQUENCE 659 AA; 70670 MW; 9C2AA1CF1EP2BAF5 CRC64;
Query Match	
Best Local Similarity 100.0%; Score 13; DB 2; Length 659;	
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	18 GGVLAAALAAAYCLS 30
Dd	559 GGVLAAALAAAYCLS 571
RESULT 54	
Q68K67_9HEPC	
ID	Q68K67_9HEPC PRELIMINARY; PRT; 659 AA.
AC	Q68K67;
DT	25-OCT-2004 (TrEMBLrel. 28, Created)
DT	25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

```
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Capid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70407 MW; 3B591F38C1521468 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 56
Q68K60_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K60;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685584; AAT94254.1; -; Genomic_RNA.
DR SMR; Q68K59; 6-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:virial transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a_.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70540 MW; C78D2F361309549B CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 58
Q68K56_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K56;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685593; AAT94257.1; -; Genomic_RNA.
```

DR SMR; Q68K56; 1-552.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70619 MW; 3A99861532AE95DE CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 559 GGVLAALAAAYCLS 571

RESULT 59
Q68K55_9HEPC
ID Q68K55_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K55;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685594; AAT94258.1; -; Genomic_RNA.
DR SMR; Q68K55; 1-552.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70619 MW; 3A99861532AE95DE CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 559 GGVLAALAAAYCLS 571

RESULT 59
Q68K55_9HEPC
ID Q68K55_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K55;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685594; AAT94258.1; -; Genomic_RNA.
DR SMR; Q68K55; 1-552.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.

FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70581 MW; 5A1A5C672BC22580 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 559 GGVLAALAAAYCLS 571

RESULT 60
Q68K54_9HEPC
ID Q68K54_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K54;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685595; AAT94259.1; -; Genomic_RNA.
DR SMR; Q68K54; 1-552.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70524 MW; 2177ESD3056081A8 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 559 GGVLAALAAAYCLS 571

RESULT 61
Q68K49_9HEPC
ID Q68K49_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K49;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685595; AAT94259.1; -; Genomic_RNA.
DR SMR; Q68K54; 1-552.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70524 MW; 2177ESD3056081A8 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
|||||
Db 559 GGVLAALAAAYCLS 571

RESULT 61
Q68K49_9HEPC
ID Q68K49_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K49;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685595; AAT94259.1; -; Genomic_RNA.
DR SMR; Q68K54; 1-552.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.

DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685608; AAT94264.1; -, Genomic_RNA.
DR SMR; Q68K49; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polypeptide;
KW Structural protein.
FT NON_TER 1
FT SEQUENCE 659 AA; 70663 MW; 6BBAD07E0C4E3B6 CRC64;
SQ
Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571
RESULT 62
Q68K47_9HEPC
ID Q68K47_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K47;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685614; AAT94266.1; -, Genomic_RNA.
DR SMR; Q68K47; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polypeptide;
KW Structural protein.
FT NON_TER 1
FT SEQUENCE 659 AA; 70663 MW; 6BBAD07E0C4E3B6 CRC64;
SQ
Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571
RESULT 62
Q68K47_9HEPC
ID Q68K47_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K47;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685614; AAT94266.1; -, Genomic_RNA.
DR SMR; Q68K47; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polypeptide;
KW Structural protein.
FT NON_TER 1
FT SEQUENCE 659 AA; 70655 MW; D3C05E533BF2CBF7 CRC64;
SQ
Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571
RESULT 63
Q68K45_9HEPC
ID Q68K45_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K45;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685617; AAT94268.1; -, Genomic_RNA.
DR SMR; Q68K45; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polypeptide;
KW Structural protein.
FT NON_TER 1
FT SEQUENCE 659 AA; 70785 MW; 47396A93B854B799 CRC64;
SQ
Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 64
Q68K43_9HEPC
ID Q68K43_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K43;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY685623; AAT94270.1; -; Genomic_RNA.
DR SMR; Q68K43; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70552 MW; 0651638FB6C80DE6 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 65
Q68K43_9HEPC
ID Q68K43_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K43;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY685623; AAT94270.1; -; Genomic_RNA.
DR SMR; Q68K43; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70714 MW; 69601C9A4B7B73AF CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 66
Q68K41_9HEPC
ID Q68K41_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K41;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY685626; AAT94272.1; -; Genomic_RNA.
DR SMR; Q68K41; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70552 MW; 0651638FB6C80DE6 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 67
Q68K41_9HEPC
ID Q68K41_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K41;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY685626; AAT94272.1; -; Genomic_RNA.
DR SMR; Q68K41; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70552 MW; 0651638FB6C80DE6 CRC64;
```

```
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70758 MW; 2AB1EFAB6020A21A CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 67
Q68K39_9HEPC
ID Q68K39_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K39;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685630; AAT94274.1; -; Genomic_RNA.
DR SMR; Q68K39; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70665 MW; EBA64CF32E067607 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 69
Q68K37_9HEPC
ID Q68K37_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K37;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685634; AAT94276.1; -; Genomic_RNA.
DR SMR; Q68K37; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70607 MW; A181B5355B1B6C53 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 559 GGVLAALAAAYCLS 571

RESULT 68
```



```
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RT Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685642; AAT94281.1; -; Genomic_RNA.
DR SMR; Q68K32; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0003676; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; DEXDC; 1.
DR SMART; SM00487; Helicase_C; 1.
DR CAPSID protein; Polyprotein; Structural protein.
KW CAPSID protein; Polyprotein; Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70664 MW; 463B7B56D85D6FB1 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 73
Q68K31_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K31;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RT Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685643; AAT94282.1; -; Genomic_RNA.
DR SMR; Q68K31; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0003676; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR SMART; SM00487; Helicase_C; 1.
DR CAPSID protein; Polyprotein; Structural protein.
KW CAPSID protein; Polyprotein; Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70670 MW; DA79D3CF2DCAFD1F CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 75
```

```
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70490 MW; F63FA0C0FCF3E90A CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 74
Q68K30_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K30;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RT Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685644; AAT94283.1; -; Genomic_RNA.
DR SMR; Q68K30; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0003676; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70670 MW; DA79D3CF2DCAFD1F CRC64;

Query Match 11.0%; Score 13; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 75
```



```

Q68K63_9HEPC
ID Q68K63_9HEPC PRELIMINARY; PRT; 660 AA.
AC Q68K63;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RA "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685572; AAT94250.1; -; Genomic_RNA.
DR SMR; Q68K63; 7-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON TER 1
FT NON TER 660
SQ SEQUENCE 660 AA; 70478 MW; DA8B21700E23599A CRC64;

Query Match 11.0%; Score 13; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 76
Q68K38_9HEPC
ID Q68K38_9HEPC PRELIMINARY; PRT; 660 AA.
AC Q68K38;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RA "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685633; AAT94275.1; -; Genomic_RNA.
DR SMR; Q68K38; 1-552.

```

```

DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON TER 1
FT NON TER 660
SQ SEQUENCE 660 AA; 70594 MW; A185CCAD99F2027C CRC64;

Query Match 11.0%; Score 13; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.001;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
DB 559 GGVLAALAAAYCLS 571

RESULT 77
Q68K29_9HEPC
ID Q68K29_9HEPC PRELIMINARY; PRT; 661 AA.
AC Q68K29;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RA "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685647; AAT94284.1; -; Genomic_RNA.
DR SMR; Q68K29; 1-554.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a_N.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
KW Capsid protein; Polyprotein; Structural protein.
FT NON TER 1
FT NON TER 661
SQ SEQUENCE 661 AA; 71409 MW; 2BED0D99BF836CF2 CRC64;

```

Query Match 11.0%; Score 13; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.001; 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0;

QY 18 GGVLAALAAAYCIS 30
|||||
DB 561 GGVLAALAAAYCIS 573

RESULT 78

Q81756_9HEPC
ID Q81756_9HEPC PRELIMINARY; PRT; 2436 AA.
AC Q81756_9HEPC PRELIMINARY; PRT; 2436 AA.
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-2004 (TRENBLrel. 26, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Choo Q.-L., Richman K., Han J.;
RL Submitted (MAY-1990) to the EMBL/GenBank/DBJ databases.
DR EMBL; M32084; AA45677.1; -; Genomic_RNA.
DR PIR; P80326; P80326.
DR PIR; P80327; P80327.
DR PIR; P80328; P80328.
DR HSP; P27958; LA1V.
DR SMR; Q81756; 579-1207, 1558-1720, 1971-2436.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptidein. 1
FT NON_TER 2436
FT SEQUENCE 2436 AA; 264737 MW; D7B9872900B3125 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 2436;
Best Local Similarity 100.0%; Pred. No. 0.0028; 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0; Indels 0;

QY 18 GGVLAALAAAYCIS 30
|||||
DB 1214 GGVLAALAAAYCIS 1226

RESULT 79

Q5UL15_9HEPC
ID Q5UL15_9HEPC PRELIMINARY; PRT; 2742 AA.
AC Q5UL15;
DT 01-FEB-2005 (TRENBLrel. 29, Created)
DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PD101;
RX PubMed=15939788; DOI=10.1084/jem.20042284;
RA Tester I., Smyk-Pearson S., Wang P., Wertheimer A., Yao E.,
RA Lewinsohn D.M., Tavis J.E., Rosen H.R.;
RT "Immune evasion versus recovery after acute hepatitis C virus
infection from a shared source";
RL J. Exp. Med. 201;1725-1731(2005).
DR EMBL; AY695437; AAV49743.1; -; Genomic DNA.
DR SMR; Q5UL15; 1029-1657, 2008-2170, 2421-2742.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptidein.
FT NON_TER 2742
FT SEQUENCE 2742 AA; 297303 MW; 38AC3AAB861E34F2 CRC64;

```
Query Match 11.0%; Score 13; DB 2; Length 2742;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 80
Q5UL16_9HEPC PRELIMINARY; PRT; 2742 AA.
AC Q5UL16;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PD102;
RX PubMed=1593788; DOI=10.1084/jem.20042284;
RA Tester I., Smyk-Pearson S., Wang P., Wertheimer A., Yao E.,
RA Lewinson D.M., Tavis J.E., Rosen H.R.;
RT "Immune evasion versus recovery after acute hepatitis C virus
RT infection from a shared source.";
RL J. Exp. Med. 201:1725-1731 (2005).
DR EMBL, AF695436; AAV49742.1; -; Genomic DNA.
DR SMR; Q5UL16; 1029-1657, 2008-2170, 2421-2742.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0008026; P: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; P: RNA binding; IEA.
DR GO; GO:0003723; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F: serine-type peptidase activity; IEA.
DR GO; GO:0008236; F: structural molecule activity; IEA.
DR GO; GO:0005198; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P: transcription; IEA.
DR GO; GO:0006350; P: viral genome replication; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.

KW Polyprotein.
FT NON_TER 2742 2742
SQ SEQUENCE 2742 AA; 297263 MW; B45ABBA437CD7891 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 2742;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1664 GGVLAALAAAYCLS 1676

RESULT 81
Q6IX04_9HEPC PRELIMINARY; PRT; 2908 AA.
AC Q6IX04;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV1a;
RA Brann T.W., Kottlilil S., Polis M., Imanichi T.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY615798; AAT44836.1; -; Genomic RNA.
DR HSSP; P27958; 1A1V.
DR SMR; Q6IX04; 1029-1657, 2008-2170, 2421-2908.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0008026; P: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: RNA binding; IEA.
DR GO; GO:0003723; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; P: structural molecule activity; IEA.
DR GO; GO:0016740; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P: transcription; IEA.
DR GO; GO:0006350; P: viral genome replication; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
```

DR Pfam: PF00998; RDRP_3; 1.
DR SMART: SMO0487; DEXDC; 1.
DR SMART: SMO0490; HELICC; 1.
KW Polyprotein.
FT NON_TER 2908 2908
SQ SEQUENCE 2908 AA; 315745 MW; BF5A4BC591498A4F CRC64;

Query Match 11.0%; Score 13; DB 2; Length 2908;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLIS 30
||| ||||| |||||
Db 1664 GGVLAALAAAYCLIS 1676

RESULT 82
POLG_HCV1
ID POLG_HCV1 STANDARD; PRT; 3010 AA.
AC P26664; Q91PES;
DT 01-AUG-1992 (Rel. 23, Created)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Genome polyprotein [contains: Core protein p19 (Capsid protein C)
DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
DE E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine
DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.-)
DE (Hepacivirin) (NS3P) (p70); Nonstructural protein 4A (NS4A) (p8);
DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
OS Hepatitis C virus (isolate 1) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11104;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=91172826; PubMed=1848704;
RA Choo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
RA Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,
RA Bradley D.W., Kuo G., Houghton M.;
RT "Genetic organization and diversity of the hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455 (1991).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RC STRAIN=Isolate infectious clone pHCV-1/SF;
RX MEDLINE=21362212; PubMed=11369872;
RA Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
RT "Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
RT sequence.";
RL J. Gen. Virol. 82:1291-1297 (2001).
RN [3]
RP PROTEIN SEQUENCE OF 1-15, AND SUBCELLULAR LOCATION OF CORE PROTEIN.
RX PubMed=7491770;
RA Lo S.-Y., Wasiar F., Hwang S.B., Lai M.M.C., Ou J.-H.;
RT "Differential subcellular localization of hepatitis C virus core gene
RT products.";
RL Virology 213:455-461 (1995).
RN [4]
RP FUNCTION OF CORE PROTEIN.
RX MEDLINE=96136654; PubMed=8533458; DOI=10.1016/0168-1702(95)00034-N;
RA Ray R.B., Lagging L.M., Meyer K., Steele R., Ray R.;
RT "Transcriptional regulation of cellular and viral promoters by the
RT hepatitis C virus core protein.";
RL Virus Res. 37:209-220 (1995).
RN [5]
RP INTERACTION OF CORE PROTEIN WITH E1 PROTEIN.
RX PubMed=8764026;
RA Lo S.-Y., Selby M.J., Ou J.-H.;
RT "Interaction between hepatitis C virus core protein and E1 envelope
RT protein.";
RL J. Virol. 70:5177-5182 (1996).
RN [6]
RP FUNCTION OF CORE PROTEIN.
RX MEDLINE=97124641; PubMed=8955036; DOI=10.1006/viro.1996.0644;
RA Ray R.B., Meyer K., Ray R.;
RT "Suppression of apoptotic cell death by hepatitis C virus core
RT protein.";
RL Virology 226:176-182 (1996).
RN [7]
RP FUNCTION OF CORE PROTEIN.
RX MEDLINE=97268985; PubMed=9110985; DOI=10.1074/jbc.272.17.10983;
RA Ray R.B., Steele R., Meyer K., Ray R.;
RT "Transcriptional repression of p53 promoter by hepatitis C virus core
RT protein.";
RL J. Biol. Chem. 272:10983-10986 (1997).
RN [8]
RP INTERACTION OF NS5A WITH HUMAN PKR.
RX MEDLINE=97288299; PubMed=9143277; DOI=10.1006/viro.1997.8493;
RA Gale M.J. Jr., Korth M.J., Tang N.M., Tan S.-L., Hopkins D.A.,
RA Dever T.E., Polyak S.J., Gretch D.R., Katze M.G.;
RT "Evidence that hepatitis C virus resistance to interferon is mediated
RT through repression of the PKR protein kinase by the nonstructural 5A
RT protein.";
RL Virology 230:217-227 (1997).
RN [9]
RP FUNCTION OF CORE PROTEIN.
RX MEDLINE=98201630; PubMed=9524287; DOI=10.1016/S0378-1119(98)00030-4;
RA Ray R.B., Steele R., Meyer K., Ray R.;
RT "Hepatitis C virus core protein represses p21WAF1/Cip1/Sid1 promoter
RT activity.";
RL Gene 208:331-336 (1998).
RN [10]
RP FUNCTION OF CORE PROTEIN.
RX PubMed=9811706;
RA Shrivastava A., Manna S.K., Ray R., Aggarwal B.B.;
RT "Ectopic expression of hepatitis C virus core protein differentially
RT regulates nuclear transcription factors.";
RL J. Virol. 72:9722-9728 (1998).
RN [11]
RP INTERACTION OF E2 WITH HUMAN PKR.
RX MEDLINE=99322192; PubMed=10390359; DOI=10.1126/science.285.5424.107;
RA Taylor D.R., Shi S.T., Romano P.R., Barber G.N., Lai M.M.C.;
RT "Inhibition of the interferon-inducible protein kinase PKR by HCV E2
RT protein.";
RL Science 285:107-110 (1999).
RN [12]
RP FUNCTION OF E2.
RX PubMed=11152499; DOI=10.1128/JVI.75.3.1265-1273.2001;
RA Taylor D.R., Tian B., Romano P.R., Hannebusch A.G., Lai M.M.C.,
RA Mathews M.B.;
RT "Hepatitis C virus envelope protein E2 does not inhibit PKR by simple
RT competition with autophosphorylation sites in the RNA-binding
RT domain.";
RL J. Virol. 75:1265-1273 (2001).
RN [13]
RP TOPOLOGY OF NS2 PROTEIN.
RX MEDLINE=22194337; PubMed=12082096; DOI=10.1074/jbc.M202304200;
RA Yamaga A.K., Ou J.-H.;
RT "Membrane topology of the hepatitis C virus NS2 protein.";
RL J. Biol. Chem. 277:33228-33234 (2002).
RN [14]
RP INTERACTION OF NS5A WITH HUMAN SRC-FAMILY KINASES, AND MUTAGENESIS OF
RP 2000-PRO-2003; 2314-PRO-2317 AND 2321-PRO-2325.
RX PubMed=14933658; DOI=10.1093/vir.0.19691-0;
RA Macdonald A., Crowder K., Street A., McCormick C., Harris M.;
RT "The hepatitis C virus NS5A protein binds to members of the Src family
RT of tyrosine kinases and regulates kinase activity.";
RL J. Gen. Virol. 85:721-729 (2004).
RN [15]
RP SUBCELLULAR LOCATION OF CORE PROTEIN.
RX PubMed=1524168; DOI=10.1128/JVI.78.15.7958-7968.2004;
RA Schwer B., Ren S., Pietschmann T., Kartenbeck J., Kaehle K.,
RA Bartenschlager R., Yen T.S.B., Ott M.;
RT "Targeting of hepatitis C virus core protein to mitochondria through a
RT novel C-terminal localization motif.";
RL J. Virol. 78:7958-7968 (2004).

[16]
RN FUNCTION OF NS5A. DOI=10.1099/vir.0.80728-0;
RX PubMed=15784895; Kallampakou K.I., Kalamvoki M., Mavromara P.;
RA "Hepatitis C virus (HCV) NS5A protein downregulates HCV IRES-dependent
RT translation.";
RL J. Gen. Virol. 86:1015-1025(2005).
RN [17]
RP REVIEW.
RX PubMed=10718937; DOI=10.1046/j.1365-2893.2000.00201.x;
RA McLauchlan J.;
RT "Properties of the hepatitis C virus core protein: a structural
RL protein that modulates cellular processes.";
RN J. Viral Hepat. 7:2-14(2000).
[18]
RP REVIEW.
RX PubMed=14752815; DOI=10.1002/hep.20032;
RA Penin F., Dublesson J., Rey F.A., Moradpour D., Pawlotsky J.-M.;
RT "Structural biology of hepatitis C virus.";
RL Hepatology 39:5-19(2004).
[19]
CC -1- FUNCTION: Core protein packages viral RNA to form a viral
CC nucleocapsid, and promotes virion budding. Modulates viral
CC translation initiation by interacting with HCV IRES and 40S
CC ribosomal subunit. Also regulates many host cellular functions
CC such as signaling pathways and apoptosis. Prevents the
CC establishment of cellular antiviral state by blocking the
CC interferon-alpha/beta (IFN-alpha/beta) and IFN-gamma signaling
CC pathways and by inducing human STAT1 degradation. Plays an
CC important role in virus-mediated cell transformation leading to
CC hepatocellular carcinomas. Interacts with, and activates STAT3
CC leading to cellular transformation. May repress the promoter of
CC p53, and sequester CREB3 and Sp110 isoform 3/Sp110b in the
CC cytoplasm. Also represses cell cycle negative regulating factor
CC CDKN1A, thereby interrupting an important check point of normal
CC cell cycle regulation. Targets transcription factors involved in
CC the regulation of inflammatory responses and in the immune
CC response: suppresses NK-kappaB activation, and activates AP-1.
CC Mediates apoptotic pathways through association with TNF-type
CC receptors TNFRSF1A and LTBR, although its effect on death
CC receptors-induced apoptosis remains controversial. Enhances TRAIL
CC mediated apoptosis, suggesting that it might play a role in
CC immune-mediated liver cell injury. Secreted core protein is able
CC to bind C1qR1 at the T-cell surface, resulting in down-regulation
CC of T-lymphocytes proliferation. May transactivate human MYC, Rous
CC sarcoma virus LTR, and SV40 promoters. May suppress the human FOS
CC and HIV-1 LTR activity. May alter lipid metabolism by interacting
CC with hepatocellular proteins involved in lipid accumulation and
CC storage.
CC -1- FUNCTION: Envelope glycoproteins E1 and E2 are involved in virus
CC attachment to the host cell as well as in virus endocytosis and
CC fusion with host membrane (by similarity). E2 inhibits human PRKR
CC activation, preventing the establishment of an antiviral state.
CC -1- FUNCTION: p7 seems to be a hexameric ion channel protein
CC (viroporin) and is inhibited by the antiviral drug amantadine.
CC Also inhibited by long-alkyl-chain iminosugar derivatives.
CC Essential for infectivity (by similarity).
CC -1- FUNCTION: Protease NS2-3, which is a putative cysteine protease,
CC is responsible for the autocatalytic cleavage of NS2-NS3. Seems to
CC undergo self-inactivation following maturation.
CC -1- FUNCTION: NS3 displays three distinct enzymatic activities: serine
CC protease, NTPase and RNA helicase. NS3 serine protease, in
CC association with NS4A, is responsible for the cleavages of NS3-
CC NS4A, NS4A-NS4B, NS4B-NS5A and NS5A-NS5B. NS3/NS4A complex also
CC prevents phosphorylation of human IRE3, thus preventing the
CC establishment of dsRNA induced antiviral state. NS3 RNA helicase
CC binds to RNA and unwinds dsRNA in the 3' to 5' direction (by
CC similarity).
CC -1- FUNCTION: NS4B may induce a specific membrane alteration that
CC serves as a scaffold for the virus replication complex. This
CC membrane alteration gives rise to the so called ER-derived
CC membranous web that contains the replication complex (by
CC similarity).

Query Match 11.0%; Score 13; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 1663 GGVLAALAAAYCLS 1675
RESULT 83
POLG_HCVH STANDARD; PRT; 3010 AA.
ID AC P27958; O36579; O36608; O36609; O36610;
DT 01-AUG-1992 (Rel. 23, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C)
DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
DE E2 (NS1) (gp68) (gp70); P7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine
DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.-)
DE (Hepacivirin) (NS3p) (p70); Nonstructural protein 4A (NS4A) (p8);
DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
OS Hepatitis C virus (isolate H) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11108;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=92052236; PubMed=1658800;
RA Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
RA Prince A.M.;
RT "Genomic structure of the human prototype strain H of hepatitis C
RT virus: comparison with American and Japanese isolates.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX STRAIN=Isolate H77;
RX MEDLINE=97373636; PubMed=9228008; DOI=10.1126/science.277.5325.570;
RA Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.M.,
RA Rice C.M.;
RT "Transmission of hepatitis C by intrahepatic inoculation with
RT transcribed RNA.";
RL Science 277:570-574(1997).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=Isolate H77;
RX MEDLINE=97385173; PubMed=9238047; DOI=10.1073/pnas.94.16.8738;
RA Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
RT "Transcripts from a single full-length cDNA clone of hepatitis C virus
RT are infectious when directly transfected into the liver of a
RT chimpanzee.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
RN [4]
RP IDENTIFICATION OF THE CYSTEINE PROTEASE, AND MUTAGENESIS OF HIS-951;
RX CVS-992 AND SER-1164.
RX PubMed=8248148;
RA Grakoui A., McCourt D.W., Wychowski C., Feinstone S.M., Rice C.M.;
RT "A second hepatitis C virus-encoded proteinase.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:10583-10587(1993).
RN [5]
RP PROCESSING OF POLYPROTEIN.
RX PubMed=7679746;
RA Grakoui A., Wychowski C., Lin C., Feinstone S.M., Rice C.M.;
RT "Expression and identification of hepatitis C virus polyprotein
RT cleavage products.";
RL J. Virol. 67:1385-1395(1993).
RN [6]
RP SUBCELLULAR LOCATION OF NS5A.
RX MEDLINE=97136712; PubMed=8982089; DOI=10.1016/S0378-1119(96)00555-0;
RA Ide Y., Zhang L., Chen M., Inchauspe G., Bahl K., Sasaguri Y.,
RA Padmanabhan R.;
RT "Characterization of the nuclear localization signal and subcellular

RT distribution of hepatitis C virus nonstructural protein NS5A.";
RL Gene 182:203-211(1996).
RN [7]
RP MUTAGENESIS OF SER-2320.
RX MEDLINE=99419094; PubMed=10488152; DOI=10.1074/jbc.274.39.28011;
RA Reed K.E., Rice C.M.;
RT "Identification of the major phosphorylation site of the hepatitis C
RT virus H strain NS5A protein as serine 2321.";
RL J. Biol. Chem. 274:28011-28018(1999).
RN [8]
RP ROLE OF TRANSMEMBRANE DOMAINS OF E1 AND E2.
RX PubMed=10729138; DOI=10.1128/JVI.74.8.3623-3633.2000;
RA Cocquerel L., Wychowski C., Minner F., Penin F., Dubuisson J.;
RT "Charged residues in the transmembrane domains of hepatitis C virus
RT glycoproteins play a major role in the processing, subcellular
RT localization, and assembly of these envelope proteins.";
RL J. Virol. 74:3623-3633(2000).
RN [9]
RP INTERACTION OF CORE PROTEIN WITH HUMAN CLOR1.
RX PubMed=11086025;
RA Kittlesen D.J., Chianese-Bullock K.A., Yao Z.Q., Braciale T.J.,
RA Hahn Y.S.;
RT "Interaction between complement receptor gC1qR and hepatitis C virus
RT core protein inhibits T-lymphocyte proliferation.";
RL J. Clin. Invest. 106:1239-1249(2000).
RN [10]
RP CHARACTERIZATION OF HVR1 REGION.
RX PubMed=11356980; DOI=10.1128/JVI.75.12.5703-5710.2001;
RA Penin F., Combet C., Germanidis G., Frainais P.-O., Deleage G.,
RA Pawlotsky J.-M.;
RT "Conservation of the conformation and positive charges of hepatitis C
RT virus E2 envelope glycoprotein hypervariable region 1 points to a role
RT in cell attachment.";
RL J. Virol. 75:5703-5710(2001).
RN [11]
RP TOPOLOGY OF NS5B.
RX MEDLINE=21570199; PubMed=11557752; DOI=10.1074/jbc.M103358200;
RA Schmidt-Mende J., Bieck E., Huegle T., Penin F., Rice C.M., Blum H.E.,
RA Moradpour D.;
RT "Determinants for membrane association of the hepatitis C virus RNA-
RT dependent RNA polymerase.";
RL J. Biol. Chem. 276:44052-44063(2001).
RN [12]
RP TOPOLOGY OF ENVELOPE GLYCOPROTEINS E1 AND E2.
RX MEDLINE=2260489; PubMed=12065403; DOI=10.1093/emboj/cdf295;
RA Cocquerel L., Op de Beeck A., Lambot M., Rousset J., Delgrange D.,
RA Pillet A., Wychowski C., Penin F., Dubuisson J.;
RT "Topological changes in the transmembrane domains of hepatitis C virus
RT envelope glycoproteins.";
RL EMBO J. 21:2893-2902(2002).
RN [13]
RP TOPOLOGY AND SUBCELLULAR LOCATION OF P7.
RX PubMed=11907211; DOI=10.1128/JVI.76.8.3720-3730.2002;
RA Carriere-Kremer S., Montpeller-Pala C., Cocquerel L., Wychowski C.,
RA Penin F., Dubuisson J.;
RT "Subcellular localization and topology of the p7 polypeptide of
RT hepatitis C virus.";
RL J. Virol. 76:3720-3730(2002).
RN [14]
RP TOPOLOGY OF NS5A.
RX MEDLINE=21864171; PubMed=11744739; DOI=10.1074/jbc.M11289200;
RA Brass V., Bieck E., Montserret R., Woelk B., Hellings J.A., Blum H.E.,
RA Penin F., Moradpour D.;
RT "An amino-terminal amphipathic alpha-helix mediates membrane
RT association of the hepatitis C virus nonstructural protein 5A.";
RL J. Biol. Chem. 277:8130-8139(2002).
RN [15]
RP REPLICATION COMPLEX.
RX PubMed=12021330; DOI=10.1128/JVI.76.12.5974-5984.2002;
RA Egger D., Woelk B., Gosert R., Bianchi L., Blum H.E., Moradpour D.,
RA Bienz K.;
RT "Expression of hepatitis C virus proteins induces distinct membrane
RT alterations including a candidate viral replication complex.";

RL J. Virol. 76:5974-5984(2002).
RN [16]
RP REPLICATION COMPLEX.
RX PubMed=12692249; DOI=10.1128/JVI.77.9.5487-5492.2003;
RA Gosert R., Egger D., Lohmann V., Bartenschlager R., Blum H.E.,
RA Bienz K., Moradpour D.;
RT "Identification of the hepatitis C virus RNA replication complex in
RT Huh-7 cells harboring subgenomic replicons.";
RL J. Virol. 77:5487-5492(2003).
RN [17]
RP INHIBITION OF P7 BY LONG-ALKYL-CHAIN IMINOSUGAR DERIVATIVES.
RX MEDLINE=22631661; PubMed=12719519; DOI=10.1073/pnas.1031527100;
RA Pavlovic D., Neville D.C., Argaud O., Blumberg B., Dwek R.A.,
RA Fischer W.B., Zitzmann N.;
RT "The hepatitis C virus p7 protein forms an ion channel that is
RT inhibited by long-alkyl-chain iminosugar derivatives.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:6104-6108(2003).
RN [18]
RP MUTAGENESIS OF LYS-778 AND ARG-780.
RX STRAIN=Isolate H77;
RA PubMed=14504405; DOI=10.1073/pnas.1834545100;
RA Sakai A., Claire M.S., Faulk K., Govindarajan S., Emerson S.U.,
RA Purcell R.H., Bukh J.;
RT "The p7 polypeptide of hepatitis C virus is critical for infectivity
RT and contains functionally important genotype-specific sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11646-11651(2003).
RN [19]
RP INTERACTIONS BETWEEN NONSTRUCTURAL PROTEINS.
RX PubMed=12692242; DOI=10.1128/JVI.77.9.5401-5414.2003;
RA Dimitrova M., Imbert I., Kieny M.P., Schuster C.;
RT "Protein-protein interactions between hepatitis C virus nonstructural
RT proteins.";
RL J. Virol. 77:5401-5414(2003).
RN [20]
RP TOPOLOGY AND SUBCELLULAR LOCATION OF NS4B.
RX STRAIN=Isolate H77;
RA PubMed=12692244; DOI=10.1128/JVI.77.9.5428-5438.2003;
RA Lundin M., Monne M., Wiedell A., Von Heijne G., Persson M.A.A.;
RT "Topology of the membrane-associated hepatitis C virus protein NS4B.";
RL J. Virol. 77:5428-5438(2003).
RN [21]
RP INTERACTION OF E1/E2 HETERODIMER WITH HUMAN CD81.
RX PubMed=12970454; DOI=10.1128/JVI.77.19.10677-10683.2003;
RA Cocquerel L., Kuo C.-C., Dubuisson J., Levy S.;
RT "CD81-dependent binding of hepatitis C virus E1E2 heterodimers.";
RL J. Virol. 77:10677-10683(2003).
RN [22]
RP INTERACTION OF E1/E2 HETERODIMER WITH HUMAN CD81.
RX MEDLINE=22503048; PubMed=12615904; DOI=10.1084/jem.20021756;
RA Bartosch B., Dubuisson J., Cosset F.-L.;
RT "Infectious hepatitis C virus pseudo-particles containing functional
RT E1-E2 envelope protein complexes.";
RL J. Exp. Med. 197:633-642(2003).
RN [23]
RP INTERACTION OF E1/E2 HETERODIMER WITH HUMAN CD81.
RX MEDLINE=22928135; PubMed=12913001; DOI=10.1074/jbc.M305289200;
RA Bartosch B., Vitelli A., Granier C., Goujon C., Dubuisson J.,
RA Pascale S., Scarselli E., Cortese R., Nicosia A., Cosset F.-L.;
RT "Cell entry of hepatitis C virus requires a set of co-receptors that
RT include the CD81 tetraspanin and the SR-B1 scavenger receptor.";
RL J. Biol. Chem. 278:41624-41630(2003).
RN [24]
RP CHARACTERIZATION OF E1 AND E2.
RX PubMed=14990718; DOI=10.1128/JVI.78.6.2994-3002.2004;
RA Op De Beeck A., Voisset C., Bartosch B., Ciczora Y., Cocquerel L.,

Query Match 11.0%; Score 13; DB 1; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAALAAAYCLS 30
DB 1663 GGVLAAALAAAYCLS 1675

RESULT 84

Q6DQ94_9HEPC PRELIMINARY; PRT; 3011 AA.
AC Q6DQ94;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Khaja M.N., Madhavi C., Habibullah C.M., Guntaka R.V.;
RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AV651061; AAT69968.1; -; Genomic_RNA.
DR SMR; Q6DQ94; 1029-1657, 2421-2950.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3011 AA; 327040 MW; DCACA9EFB33D218 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

Db 1664 GGVLAALAAAYCLS 1676

RESULT 85

Q913D4_9HEPC PRELIMINARY; PRT; 3011 AA.
AC Q913D4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Guntaka R.V., Munnally S.K., Khaja M.N., Kota K.K., Ramana V.K.,
RA Swamyathan S.K., Sakata Y., Habeebullah C.M.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY051292; AAK95832.1; -; mRNA.
DR HSSP; Q913D4; 1029-1657, 2421-2950.
DR SMR; Q913D4; 1029-1657, 2421-2950.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3011 AA; 327235 MW; 57A21964B4227B60 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30

|||||

Db 1664 GGVLAALAAAYCLS 1676

RESULT 86

Q9DIT6_9HEPC PRELIMINARY; PRT; 3011 AA.
ID Q9DIT6_9HEPC

AC Q9DIT6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Genomic RNA for polyprotein gene.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21014672; PubMed=11115058;
RX DOI=10.1046/j.1365-2893.2000.00259.x;
RA Kumar U., Tuthill T., Thomas H.C., Monjardino J.;
RT "Sequence, expression and reconstitution of an HCV genome from a
RT British isolate derived from a single blood donation.";
RL J. Viral Hepat. 7:459-465(2000).
DR EMBL; AJ278830; CAC03609.1; -; Genomic_RNA.
DR PIR; PS0326; PS0327.
DR PIR; PS0327; PS0328.
DR PIR; PS0328; PS0328.
DR HSP; P27958; 1AIV.
DR SMR; Q9DIT6; 1029-1657, 2421-2950.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0002236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR011492; Flavi_DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein. 2 191 core protein.
FT CHAIN 192 383 envelop protein 1.
FT CHAIN 384 809 envelop protein 2.
FT CHAIN 810 1026 non-structural protein 2.
FT CHAIN 1027 1657 non-structural protein 3.
FT CHAIN 1658 1711 non-structural protein 4a.
FT CHAIN 1712 1972 non-structural protein 4b.
FT CHAIN 1973 2420 non-structural protein 5a.

FT CHAIN 2421 3011 non-structural protein 5b.
SQ SEQUENCE 3011 AA; 327406 MW; 7B6264A74A5452D3 CRC64;
Query Match 11.0%; Score 13; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred.No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
DB 1664 GGVLAALAAAYCLS 1676
RESULT 87
Q03463_9HSPC
ID Q03463_9HSPC PRELIMINARY; PRT; 3011 AA.
AC Q03463;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044440; PubMed=1658196;
RX Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H., Machida A.,
RA Miyakawa Y., Mayumi M.;
RT "Nucleotide sequence of the genomic RNA of hepatitis C virus isolated
RT from a human carrier: comparison with reported isolates for conserved
RT and divergent regions.";
RL J. Gen. Virol. 72:2697-2704(1991).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93117120; PubMed=1335573;
RX Okamoto H., Kanai N., Mishiro S.;
RT "Full-length nucleotide sequence of a Japanese hepatitis C virus
RT isolate (HC-J1) with high homology to USA isolates.";
RL Nucleic Acids Res. 20:6410-6410(1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=91013116; PubMed=2170712;
RX Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
RA Yoshizawa H., Tsuda F., Miyakawa Y., Mayumi M.;
RT "The 5'-terminal sequence of the hepatitis C virus genome.";
RL Jpn. J. Exp. Med. 60:167-177(1990).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94174722; PubMed=7510436;
RX Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
RA Inchausti G.;
RT "Characterization and mapping of a B-cell immunogenic domain in
RT hepatitis C virus E2 glycoprotein using a yeast peptide library.";
RL Virology 200:246-255(1994).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Okamoto H.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; D10749; BAA01582.1; -; Genomic_RNA.
DR PIR; PS0326; PS0326.
DR PIR; PS0327; PS0327.
DR PIR; PS0328; PS0328.
DR PIR; PS0328; PS0328.
DR HSP; P26664; 1HEI.
DR SMR; Q03463; 1029-1657, 2421-2950.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0002236; F:serine-type peptidase activity; IEA.

DR GO: GO:0005198; P:structural molecule activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR GO: GO:0019079; P:Viral genome replication; IEA.
 DR GO: GO:0019087; P:Viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV core.
 DR InterPro: IPR002519; HCV env.
 DR InterPro: IPR002531; HCV NS1.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001490; HCV NS4b.
 DR InterPro: IPR002868; HCV NS5a.
 DR InterPro: IPR002166; HCV RdRp.
 DR InterPro: IPR001650; Helicase C.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV core; 1.
 DR Pfam: PF01539; HCV env; 1.
 DR Pfam: PF01560; HCV NS1; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF02907; HCV NS2; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF01001; HCV NS4a; 1.
 DR Pfam: PF01506; HCV NS4b; 1.
 DR Pfam: PF00271; Helicase C; 1.
 DR Pfam: PF00998; RdRp 3; 1.
 DR SMART; SM00487; DEXDc; 1.
 KW Polyprotein.
 SQ SEQUENCE 3011 AA; 327114 MW; 97E9052C0250463B CRC64;

Query Match 11.0%; Score 13; DB 2; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.0034;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
 |||||
 DB 1664 GGVLAALAAAYCLS 1676

RESULT 88
 Q81754_9HEPC
 ID Q81754_9HEPC PRELIMINARY; PRT; 3011 AA.
 AC Q81754;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HC-G9;
 RX MEDLINE=94172337; PubMed=8126459;
 RA Okamoto H., Kojima M., Sakamoto M., Iizuka H., Hadiwandowo S.,
 RA Suwignyo S., Miyakawa Y., Mayumi M.;
 RT "The entire nucleotide sequence and classification of a hepatitis C
 RT virus isolate of a novel genotype from an Indonesian patient with
 RT chronic liver disease.";
 RL J. Gen. Virol. 75:629-635 (1994).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93224896; PubMed=8385694;
 RA Simmonds P., McOmish P., Yap P.L., Chan S.-W., Lin C.K.,
 RA Dushenko G., Saeed A.A., Holmes E.C.;
 RT "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence

RT diversity.";
 RL J. Gen. Virol. 74:661-668 (1993).
 DR EMBL; D14853; BAA03581.1; -; Genomic_RNA.
 DR PIR; PQ0804; PQ0804.
 DR HSSP; Q81755; 1DXP.
 DR SMR; Q81754; 1029-1657, 2008-2170, 2421-2950.
 DR GO: GO:0019028; C:viral capsid; IEA.
 DR GO: GO:0019031; C:viral envelope; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR GO: GO:0019079; P:Viral genome replication; IEA.
 DR GO: GO:0019087; P:Viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV core.
 DR InterPro: IPR002519; HCV env.
 DR InterPro: IPR002531; HCV NS1.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001490; HCV NS4b.
 DR InterPro: IPR002868; HCV NS5a.
 DR InterPro: IPR002166; HCV RdRp.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV core; 1.
 DR Pfam: PF01539; HCV env; 1.
 DR Pfam: PF01560; HCV NS1; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF02907; HCV NS2; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF01001; HCV NS4a; 1.
 DR Pfam: PF01506; HCV NS4b; 1.
 DR Pfam: PF00998; RdRp 3; 1.
 DR SMART; SM00487; DEXDc; 1.
 KW Polyprotein.
 FT CHAIN 1 191 core.
 FT CHAIN 192 383 E1.
 FT CHAIN 384 809 E2/NS1.
 FT CHAIN 810 1006 NS2.
 FT CHAIN 1007 1657 NS3.
 FT CHAIN 1658 1972 NS4.
 FT CHAIN 1973 3011 NS5.
 SQ SEQUENCE 3011 AA; 327216 MW; 9C16C120F4E79268 CRC64;

Query Match 11.0%; Score 13; DB 2; Length 3011;
 Best Local Similarity 100.0%; Pred. No. 0.0034;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
 |||||
 DB 1664 GGVLAALAAAYCLS 1676

RESULT 89
 Q9ELS8_9HEPC
 ID Q9ELS8_9HEPC PRELIMINARY; PRT; 3011 AA.
 AC Q9ELS8;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.

OX NCBI_TaxID=11103;
RN [1] NUCLEOTIDE SEQUENCE.
RA Desai S.M., Devare S., Yamaguchi J.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed:1373489;
RA Ching W.M., Wychowski C., Beach M.J., Wang H., Davies C.L., Carl M.,
Bradley D.W., Alter H.J., Feinstein S.M., Shih J.W.;
RT "Interaction of immune sera with synthetic peptides corresponding to
the structural protein region of hepatitis C virus";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3190-3194 (1992).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE:93224886; PubMed:8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
virus: identification of a new virus type and restrictions on sequence
diversity";
RL J. Gen. Virol. 74:661-668 (1993).
DR EMBL; AF290976; AAG02099.1; -; Genomic_DNA.
DR PIR; A44150; A44150.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0326; PS0326.
DR PIR; PS0327; PS0327.
DR PIR; PS0328; PS0328.
DR HSSP; P26664; 1HEI.
DR SMR; Q9ELS8; 1029-1657, 2008-2170, 2421-2950.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR002518; Peptidase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3011 AA; 327111 MW; A6BECF5A3B3BE13F CRC64;

Query Match 11.0%; Score 13; DB 2; Length 3011;

Best Local Similarity 100.0%; Pred. No. 0.0034;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCLS 30
||| ||||| |||||
DB 1664 GGVLAALAAAYCLS 1676
RESULT 90
Q9PMU9_922ZZ PRELIMINARY; PRT; 3015 AA.
AC Q9PMU9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
DE Polyprotein.
OS synthetic construct.
OC other sequences; artificial sequences.
OX NCBI_TaxID=32630;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE:99420396; PubMed:10489358; DOI=10.1006/viro.1999.9889;
RA Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
RT "Hepatitis C virus: an infectious molecular clone of a second major
genotype (2a) and lack of viability of intertypic 1a and 2a
chimeras";
RL Virology 262:250-263 (1999).
DR EMBL; AF177039; AAF01181.1; -; Genomic_RNA.
DR EMBL; AF177037; AAF01179.1; -; Genomic_RNA.
DR PIR; PS0326; PS0326.
DR PIR; PS0327; PS0327.
DR PIR; PS0328; PS0328.
DR HSSP; P27958; 1HEI.
DR SMR; Q9PMU9; 1033-1661, 2425-2954.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Peptidase_S29.
DR InterPro; IPR002129; Pyridoxal_deC.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.

```
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.
KW Polyprotein.
SQ SEQUENCE 3015 AA; 328084 MW; E309F6318067D6CD CRC64;

  Query Match      11.0%; Score 13; DB 2; Length 3015;
  Best Local Similarity 100.0%; Pred. No. 0.0034;
  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1668 GGVLAALAAAYCLS 1680

RESULT 91
Q9PMX5 92ZZZ PRELIMINARY; PRT; 3015 AA.
AC Q9PMX5;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE Polyprotein.
OS synthetic construct.
OC other sequences; artificial sequences.
OX NCBI_TaxID=32630;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99420396; PubMed=10489358; DOI=10.1006/viro.1999.9889;
RA Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
RT "Hepatitis C virus: an infectious molecular clone of a second major
RT genotype (2a) and lack of viability of intertypic 1a and 2a
RT chimeras.";
RL Virology 262:250-263(1999).
DR EMBL; AF17040; AAF01182.1; -; Genomic RNA.
DR EMBL; AF17038; AAF01180.1; -; Genomic RNA.
DR PIR; PS0326; PS0326.
DR PIR; PS0327; PS0327.
DR PIR; PS0328; PS0328.
DR HSSP; P27958; 1HEI.
DR SMR; Q9PMX5; 1033-1661, 2425-2954.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR002129; Pyridoxal dec.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
```

```
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; UNKNOWN_1.
KW Polyprotein.
SQ SEQUENCE 3015 AA; 328159 MW; B7D23BC1F190663A CRC64;

  Query Match      11.0%; Score 13; DB 2; Length 3015;
  Best Local Similarity 100.0%; Pred. No. 0.0034;
  Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 30
Db 1668 GGVLAALAAAYCLS 1680

RESULT 92
Q68204 9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68204;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Namesn4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ib;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14240; AAC53929.1; -; Genomic RNA.
DR HSSP; P26663; 1CU1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR NON_TER 1
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15155 MW; 09D3E469C07958C3 CRC64;

  Query Match      10.2%; Score 12; DB 2; Length 138;
  Best Local Similarity 100.0%; Pred. No. 0.0027;
  Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCLS 29
Db 38 GGVLAALAAAYCLS 49

RESULT 93
Q68205 9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68205;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Namesn4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
```

```

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=1b;
RC MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14241; AAC53930.1; -; Genomic_RNA.
DR HSP; P27958; 1HEI.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15205 MW; 6D376B2DB6EADAA CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 94
Q68206_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68206;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14242; AAC53931.1; -; Genomic_RNA.
DR HSP; P27958; 1HEI.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15117 MW; 5FBC51A1B74DE13E CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 95
Q68207_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68207;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
```

```

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14243; AAC53932.1; -; Genomic_RNA.
DR HSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15173 MW; 45A8E7DD60F711D3 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 96
Q68208_9HEPC PRELIMINARY; PRT; 138 AA.
ID Q68208_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68208;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14244; AAC53933.1; -; Genomic_RNA.
DR HSP; P26663; 1CUL.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15161 MW; 98A4B2A623462898 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49
```

```
RESULT 97
Q68209_9HEPC
ID Q68209_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68209;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14245; AAC53934.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15149 MW; DBA862A0FE9E2D57 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 38 GGVLAALAAAYCL 49

RESULT 98
Q68210_9HEPC
ID Q68210_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68210;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14246; AAC53935.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15049 MW; P2E9A133AB4958F9 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
```

```
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 38 GGVLAALAAAYCL 49

RESULT 99
Q68211_9HEPC
ID Q68211_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68211;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14247; AAC53936.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15101 MW; AF97AB5576594908 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 38 GGVLAALAAAYCL 49

RESULT 100
Q68212_9HEPC
ID Q68212_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68212;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14248; AAC53937.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
```

```
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15092 MW; DD3CE09283403052 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 101
Q68213_9HEPC
ID Q68213_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68213;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14249; AAC53938.1; -; Genomic_RNA.
DR HSP; P26663; 1CUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15136 MW; 332E8B13B3592BE9 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 102
Q68215_9HEPC
ID Q68215_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68215;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
```

```
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14251; AAC53940.1; -; Genomic_RNA.
DR HSP; P26663; 1CUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15133 MW; 692DB2A62346288A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 103
Q68216_9HEPC
ID Q68216_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68216;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14252; AAC53941.1; -; Genomic_RNA.
DR HSP; P26663; 1CUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15081 MW; 7ED533A7D169FB1A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db |||||
38 GGVLAALAAAYCL 49

RESULT 104
Q68217_9HEPC
ID Q68217_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68217;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
```

```
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RA MEDLINE=95146953; PubMed=7844535;
RX Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14253; AAC53942.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15130 MW; 2AF1E92DDC7B741D CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 105
Q68218_9HEPC
ID Q68218_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68218;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RA MEDLINE=95146953; PubMed=7844535;
RX Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14254; AAC53943.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15189 MW; DB78E92DDC67040P CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 106
Q68221_9HEPC
ID Q68221_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68221;
```

```
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RA MEDLINE=95146953; PubMed=7844535;
RX Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14257; AAC53946.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15046 MW; AB9551B4B1472B9A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 107
Q68227_9HEPC
ID Q68227_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68227;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RA MEDLINE=95146953; PubMed=7844535;
RX Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14263; AAC53952.1; -; Genomic_RNA.
DR HSSP; P26663; 8OEH.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15074 MW; B2744939F0E81F28 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49
```

Db 38 GGVLAALAAAYCL 49

RESULT 108

Q68228_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68228_9HEPC PRELIMINARY; PRT; 138 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia."
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14264; AAC53953.1; -; Genomic_RNA.
DR HSP; P26663; ICUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15090 MW; BF845639F0E80028 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

Db 38 GGVLAALAAAYCL 49

RESULT 109

Q68229_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68229_9HEPC PRELIMINARY; PRT; 138 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia."
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14265; AAC53954.1; -; Genomic_RNA.
DR HSP; P26663; ICUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15110 MW; 3941892EE34FDA79 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

Db 38 GGVLAALAAAYCL 49

RESULT 110

Q68230_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68230_9HEPC PRELIMINARY; PRT; 138 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia."
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14266; AAC53955.1; -; Genomic_RNA.
DR HSP; P26663; ICUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15076 MW; 2C202A332C4A5BE9 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

Db 38 GGVLAALAAAYCL 49

RESULT 111

Q68231_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68231_9HEPC PRELIMINARY; PRT; 138 AA.
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia."
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14267; AAC53956.1; -; Genomic_RNA.
DR HSP; P27958; 1AIR.


```
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15051 MW; 643PFFP7B94CACACIE CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 112
Q68232_9HEPC
ID Q68232_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68232;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14268; AAC53957.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15087 MW; 643PFFP7B7412CFE CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 113
Q68235_9HEPC
ID Q68235_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68235;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14268; AAC53957.1; -; Genomic_RNA.
DR HSSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15087 MW; 643PFFP7B7412CFE CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 114
Q68237_9HEPC
ID Q68237_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68237;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14273; AAC53962.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15104 MW; 74E2FA08BC55391A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 115
Q68238_9HEPC
ID Q68238_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68238;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
```

```
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14271; AAC53960.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15104 MW; 585DC5A627D0F3E3 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 114
Q68237_9HEPC
ID Q68237_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68237;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215 (1995).
DR EMBL; U14273; AAC53962.1; -; Genomic_RNA.
DR HSSP; P26663; 1CUI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15104 MW; 74E2FA08BC55391A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 115
Q68238_9HEPC
ID Q68238_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68238;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
```

```
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14274; AAC53963.1; -; Genomic_RNA.
DR HSP; P26663; ICUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15173 MW; 4110A9B33E768F1B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 38 GGVLAALAAAYCL 49

RESULT 116
Q68240_9HEPC
ID Q68240_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68240;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14276; AAC53965.1; -; Genomic_RNA.
DR HSP; P27958; LAIR.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15115 MW; 6A042677B354CA7A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 38 GGVLAALAAAYCL 49

RESULT 117
Q68242_9HEPC
```

```
ID Q68242_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68242;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14278; AAC53967.1; -; Genomic_RNA.
DR HSP; P26663; ICUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15281 MW; CD5B5B3834C6070D CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 38 GGVLAALAAAYCL 49

RESULT 118
Q68244_9HEPC
ID Q68244_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68244;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1b;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14280; AAC53969.1; -; Genomic_RNA.
DR HSP; P26663; ICUL.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1
FT NON_TER 138
SQ SEQUENCE 138 AA; 15118 MW; B7F7EB2733770408 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 18 GGVLAALAAAYCL 29
DB 38 GGVLAALAAAYCL 49

RESULT 119

Q81574_9HEPC

ID Q81574_9HEPC PRELIMINARY; PRT; 172 AA.

AC Q81574;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Nonstructural protein (Fragment).

GN Name=NS3-NS4;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M86767; AAA45617.1; -; Genomic_RNA.

DR HSP; P26663; 1CU1.

DR SMR; Q81574; 1-86.

DR InterPro; IPR000745; HCV NS4a.

DR Pfam; PF01006; HCV NS4a; 1.

FT NON_TER 172

SQ SEQUENCE 172 AA; 18941 MW; 08DC37A566CD84D6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;

Best Local Similarity 100.0%; Pred. No. 0.0033;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

DB 93 GGVLAALAAAYCL 104

RESULT 120

Q81575_9HEPC

ID Q81575_9HEPC PRELIMINARY; PRT; 172 AA.

AC Q81575;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Nonstructural protein (Fragment).

GN Name=NS3-NS4;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M86768; AAA45618.1; -; Genomic_RNA.

DR HSP; P26663; 1CU1.

DR SMR; Q81575; 1-86.

DR InterPro; IPR000745; HCV NS4a.

DR Pfam; PF01006; HCV NS4a; 1.

FT NON_TER 172

SQ SEQUENCE 172 AA; 18941 MW; 3E3784DCBBF1BED CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;

Best Local Similarity 100.0%; Pred. No. 0.0033;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

DB 93 GGVLAALAAAYCL 104

RESULT 121

Q81577_9HEPC

ID Q81577_9HEPC PRELIMINARY; PRT; 172 AA.

AC Q81577;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Nonstructural protein (Fragment).

GN Name=NS3-NS4;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M86770; AAA45620.1; -; Genomic_RNA.

DR HSP; P26663; 1CU1.

DR SMR; Q81577; 1-86.

DR InterPro; IPR000745; HCV NS4a.

DR Pfam; PF01006; HCV NS4a; 1.

FT NON_TER 172

SQ SEQUENCE 172 AA; 18908 MW; 3550C722AC3E5EE9 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;

Best Local Similarity 100.0%; Pred. No. 0.0033;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

DB 93 GGVLAALAAAYCL 104

RESULT 122

Q81578_9HEPC

ID Q81578_9HEPC PRELIMINARY; PRT; 172 AA.

AC Q81578;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE Nonstructural protein (Fragment).

GN Name=NS3-NS4;

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;

RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M86772; AAA45621.1; -; Genomic_RNA.

DR HSP; P26663; 1CU1.

DR SMR; Q81578; 1-86.

DR InterPro; IPR000745; HCV NS4a.

DR Pfam; PF01006; HCV NS4a; 1.

FT NON_TER 172

SQ SEQUENCE 172 AA; 18941 MW; 08DC37A566CD84D6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;

Best Local Similarity 100.0%; Pred. No. 0.0033;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

DB 93 GGVLAALAAAYCL 104

RESULT 123

Q81579_9HEPC

ID Q81579_9HEPC PRELIMINARY; PRT; 172 AA.

```

AC Q81579;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M86773; AAA45622.1; -, Genomic_RNA.
DR HSSP; P26663; 1CUL.
DR SMR; Q81579; 1-86.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 172
SQ SEQUENCE 172 AA; 18929 MW; 237936051FCE2AD1 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 93 GGVLAALAAAYCL 104

RESULT 124
Q81580_9HEPC
ID Q81580_9HEPC PRELIMINARY; PRT; 172 AA.
AC Q81580;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M86774; AAA45623.1; -, Genomic_RNA.
DR HSSP; P26663; 1CUL.
DR SMR; Q81580; 1-86.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 172
SQ SEQUENCE 172 AA; 18935 MW; 8D1978F6E3F26336 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 93 GGVLAALAAAYCL 104

RESULT 125
Q81581_9HEPC
ID Q81581_9HEPC PRELIMINARY; PRT; 172 AA.
AC Q81581;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).

```

```

GN Name=NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M86775; AAA45624.1; -, Genomic_RNA.
DR HSSP; P26663; 1CUL.
DR SMR; Q81581; 1-86.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 172
SQ SEQUENCE 172 AA; 18941 MW; 08DC37AE66CD84D6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 93 GGVLAALAAAYCL 104

RESULT 126
Q81582_9HEPC
ID Q81582_9HEPC PRELIMINARY; PRT; 172 AA.
AC Q81582;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M86776; AAA45625.1; -, Genomic_RNA.
DR HSSP; P26663; 1CUL.
DR SMR; Q81582; 1-86.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 172
SQ SEQUENCE 172 AA; 18941 MW; 08DC37AE66CD84D6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 93 GGVLAALAAAYCL 104

RESULT 127
Q81583_9HEPC
ID Q81583_9HEPC PRELIMINARY; PRT; 172 AA.
AC Q81583;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;

```

RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M86777; AAA45626.1; -; Genomic_RNA.
DR HSSP; P26663; 1CU1.
DR SMR; Q81583; 1-86.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 18943 MW; 49D8356DC338179E CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAALAAAYCL 29
|||||
Db 93 GGVLAAALAAAYCL 104

RESULT 128
Q81584_9HEPC
ID Q81584_9HEPC PRELIMINARY; PRT; 172 AA.
AC Q81584;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name:NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL; M86778; AAA45627.1; -; Genomic_RNA.
DR HSSP; P26663; 1CU1.
DR SMR; Q81584; 1-86.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 172 172
SQ SEQUENCE 172 AA; 18941 MW; 08DC37AE66CD84D6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 0.0033;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAALAAAYCL 29
|||||
Db 93 GGVLAAALAAAYCL 104

RESULT 129
Q81573_9HEPC
ID Q81573_9HEPC PRELIMINARY; PRT; 271 AA.
AC Q81573;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name:NS3-NS4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Miller H.M., Goesser T., Pfaff E., Theilmann L.;
RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.

DR EMBL; M86766; AAA45616.1; -; Genomic_RNA.
DR HSSP; P26663; 1CU1.
DR SMR; Q81573; 2-130.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 271 271
SQ SEQUENCE 271 AA; 29796 MW; 72BEBC54E6877CD4 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 271;
Best Local Similarity 100.0%; Pred. No. 0.0047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAALAAAYCL 29
|||||
Db 137 GGVLAAALAAAYCL 148

RESULT 130
Q68K53_9HEPC
ID Q68K53_9HEPC PRELIMINARY; PRT; 652 AA.
AC Q68K53;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects."
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV685596; AAT94260.1; -; Genomic_RNA.
DR SMR; Q68K53; 1-549.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02507; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDc; 1.
KW Capsid protein; Polyprotein; Structural protein.
FT NON_TER 1 1
FT NON_TER 652 652
SQ SEQUENCE 652 AA; 69700 MW; 21B5CB69FD3C43F0 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 652;
Best Local Similarity 100.0%; Pred. No. 0.0094;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAALAAAYCL 29
|||||
Db 556 GGVLAAALAAAYCL 567

RESULT 131
Q68K69_9HEPC
ID Q68K69_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K69;

```

DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685559; AAT94244.1; -; Genomic_RNA.
DR SMR; Q68K57; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF0271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70594 MW; CFF7E6C7E0242545 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 659;
Best Local Similarity 100.0%; Pred.No. 0.0095;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 559 GGVLAALAAAYCL 570

RESULT 132
Q68K57_9HEPC
ID Q68K57_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K57;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685559; AAT94256.1; -; Genomic_RNA.
DR SMR; Q68K57; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.

```

```

DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF0271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70412 MW; 2FCC6D4A67324E1 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 659;
Best Local Similarity 100.0%; Pred.No. 0.0095;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 559 GGVLAALAAAYCL 570

RESULT 133
Q68K50_9HEPC
ID Q68K50_9HEPC PRELIMINARY; PRT; 659 AA.
AC Q68K50;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685606; AAT94263.1; -; Genomic_RNA.
DR SMR; Q68K50; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF0271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659

```



```

DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 659
SQ SEQUENCE 659 AA; 70412 MW; 2FCC6D44A67324E1 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 559 GGVLAALAAAYCL 570
|||||
|||||

RESULT 137
Q68K46_9HEPC
ID Q68K46_9HEPC PRELIMINARY; PRT; 660 AA.
AC Q68K46;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;
RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
RT HCV/HIV Coinfected Subjects.";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY685615; AAT94267.1; -; Genomic_RNA.
DR SMR; Q68K46; 1-552.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Capsid protein; Helicase; Hydrolase; Polyprotein;
KW Structural protein.
FT NON_TER 1
FT NON_TER 661
SQ SEQUENCE 661 AA; 70683 MW; 6CB5B1B0442AEB05 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 661;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 561 GGVLAALAAAYCL 572
|||||
|||||

RESULT 139
Q68K52_9HEPC
ID Q68K52_9HEPC PRELIMINARY; PRT; 661 AA.
AC Q68K52;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Parker M., Lamson D., Wroblewski D., Reilly A., Philpott S.,
RA Kleiner D., Holman S., Augenbraun M., Taylor J.;

```


RT "Sequence Analysis of Hepatitis C Virus Replication Functions in
 RT HCV/HIV Coinfected Subjects.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY685597; AAT94261.1; -; Genomic_RNA.

DR SMR; Q68K52; 1-554.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003676; F:nucleic acid binding; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000745; HCV_NS4a.
 DR InterPro; IPR004109; Peptidase_S29.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW Capsid protein; Polyprotein; Structural protein.
 FT NON TER 1 1
 FT NON TER 661 661
 SQ SEQUENCE 661 AA; 70702 MW; 1F22BF3AA974A315 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 661;
 Best Local Similarity 100.0%; Pred. No. 0.0095;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 561 GGVLAALAAAYCL 572

RESULT 140

O41809_9HEPC
 ID O41809_9HEPC PRELIMINARY; PRT; 1805 AA.
 AC O41809;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97032798; PubMed=8878547; DOI=10.1006/bbrc.1996.1540;
 RA Yeh C.T., Chu C.M., Liaw Y.F.;
 RT "Distinct composition of viral quasiespecies between ascites and serum
 RT samples from patients with late stage chronic hepatitis C.";
 RL Biochem. Biophys. Res. Commun. 227:524-529(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98033184; PubMed=937361;
 RA Yeh C.T., Lu S.C., Chu C.M., Liaw Y.F.;
 RT "Molecular cloning of a defective hepatitis C virus genome from the
 RT ascitic fluid of a patient with hepatocellular carcinoma.";
 RL J. Gen. Virol. 78:0-0(10).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Yeh C.T., Chu C.M., Lu S.J., Liaw Y.F.;
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U89019; AAB8251.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR HSSP; P26663; IQUV.
 DR SMR; O41809; 13-454, 1218-1744.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006350; P:transcription; IEA.

DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000745; HCV_NS4a.
 DR InterPro; IPR001490; HCV_NS4b.
 DR InterPro; IPR002868; HCV_NS5a.
 DR InterPro; IPR002166; HCV_RDRP.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00398; RDRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW Polyprotein.
 SQ SEQUENCE 1805 AA; 195957 MW; 28D2FD07D0F5074B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 1805;
 Best Local Similarity 100.0%; Pred. No. 0.021;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
 |||||
 DB 463 GGVLAALAAAYCL 474

RESULT 141

Q7T4V8_9HEPC
 ID Q7T4V8_9HEPC PRELIMINARY; PRT; 1984 AA.
 AC Q7T4V8;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE NS protein (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22694192; PubMed=12810084; DOI=10.1016/S0006-291X(03)01047-7;
 RA Kato N., Sugiyama K., Namba K., Dansako H., Nakamura T., Takami M.,
 RA Naka K., Nozaki A., Shimotohno K.;
 RT "Establishment of a hepatitis C virus subgenomic replicon derived from
 RT human hepatocytes infected in vitro.";
 RL Biochem. Biophys. Res. Commun. 306:756-766(2003).
 DR EMBL; AB109543; BAC77767.1; -; Genomic_RNA.
 DR HSSP; Q81755; 1DXP.
 DR SMR; Q7T4V8; 3-631, 982-1144, 1394-1923.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral genome reformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; HCV_NS4a.
 DR InterPro; IPR000745; HCV_NS4b.
 DR InterPro; IPR001490; HCV_NS5a.
 DR InterPro; IPR002868; HCV_RDRP.
 DR InterPro; IPR002166; Helicase_C.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.

DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RDRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW Capsid protein; Structural protein; Transmembrane.
 FT NON TER 1
 SQ SEQUENCE 1984 AA; 214361 MW; 8AA8198D2C7B291C CRC64;
 Query Match 10.2%; Score 12; DB 2; Length 1984;
 Best Local Similarity 100.0%; Pred. No. 0.023;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAALAAAYCL 29
 Db 638 GGVLAALAAAYCL 649
 RESULT 142
 Q81817_9HEPC PRELIMINARY; PRT; 2284 AA.
 AC Q81817;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Polypeptide.
 OS Hepatitis C virus type 2.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=40271;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=94068484; PubMed=7504283;
 RA Hijikata M., Mizushima H., Tanji Y., Komada Y., Hirowatari Y.,
 RA Akagi T., Kimura K., Shimotohno K.;
 RT "Proteolytic processing and membrane association of putative
 RT nonstructural proteins of hepatitis C virus";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10773-10777(1993).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=94333810; PubMed=8056334; DOI=10.1016/0378-1119(94)90008-6;
 RA Tanji Y., Hijikata M., Hirowatari Y., Shimotohno K.;
 RT "Identification of the domain required for trans-cleavage activity of
 RT hepatitis C viral serine proteinase";
 RL Gene 145:215-219(1994).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=95056078; PubMed=7966638;
 RA Tanji Y., Hijikata M., Hirowatari Y., Shimotohno K.;
 RT "Hepatitis C virus polypeptide processing: kinetics and mutagenic
 RT analysis of serine proteinase-dependent cleavage";
 RL J. Virol. 68:8418-8422(1994).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=95156583; PubMed=7853491;
 RA Tanji Y., Hijikata M., Satoh S., Kaneko T., Shimotohno K.;
 RT "Hepatitis C virus-encoded nonstructural protein NS4A has versatile
 RT functions in viral protein processing";
 RL J. Virol. 69:1575-1581(1995).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity";
 RL J. Gen. Virol. 72:2805-2809(1991).
 DR EMBL; D16435; BA03905.1; -; Genomic_RNA.
 DR FIR; A61196; A61196.
 DR FIR; PQ0246; PQ0246.
 DR FIR; PQ0329; PQ0329.
 DR HSSP; Q81755; 1DXP.
 DR SMR; Q81817; 303-931, 1282-1444, 1694-2223.

DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; P:serine-type peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RDRP.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR002518; Pept U39 HCV NS2.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RDRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW Polypeptide.
 FT CHAIN 21 54 Potential.
 FT CHAIN 84 300 NS2.
 FT CHAIN 301 931 NS3.
 FT CHAIN 932 985 NS4A.
 FT CHAIN 986 1246 NS4B.
 FT CHAIN 1247 1693 NS5A.
 FT CHAIN 1694 2284 NS5B.
 SQ SEQUENCE 2284 AA; 247216 MW; DC272A1517046337 CRC64;
 Query Match 10.2%; Score 12; DB 2; Length 2284;
 Best Local Similarity 100.0%; Pred. No. 0.026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 18 GGVLAALAAAYCL 29
 Db 938 GGVLAALAAAYCL 949
 RESULT 143
 Q92973_9HEPC PRELIMINARY; PRT; 2864 AA.
 AC Q92973;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HC-J4;
 RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
 RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
 RA Bukh J.;
 RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
 RT infectious in vivo";
 RL Virology 244:161-172(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;

RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054253; AAC15727.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR SNR; O92973; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferrase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR002518; Pept_U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310469 MW; 916DDA2FD0449C98 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
|||||

RESULT 144
O92974_9HEPC PRELIMINARY; PRT; 2864 AA.
AC O92974;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;

RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581789; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Ohnishi M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054257; AAC15730.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR SNR; O92974; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferrase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR002518; Pept_U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310485 MW; 32CF23E5E59C4E59 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
|||||

```
RESULT 145
O92975 9HEPC
ID O92975_9HEPC PRELIMINARY; PRT; 2864 AA.
AC O92975;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RA "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054259; AAC15731.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR SMR; O92975; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_core.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_core.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR007094; RNA_pol_PSvir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_NSI; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Pfam; PF01538; HCV_NSI; 1.
DR Pfam; PF02907; HCV_NSI; 1.
DR Pfam; PF01006; HCV_NSI; 1.
DR Pfam; PF01001; HCV_NSI; 1.

DR Pfam; PF01506; HCV_NSI; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
FT NON_TER 2864 2864
SQ SEQUENCE 2864 AA; 310419 MW; C0CD3933ED07C6A5 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 1664 GGVLAALAAAYCL 1675

RESULT 146
O92976 9HEPC
ID O92976_9HEPC PRELIMINARY; PRT; 2864 AA.
AC O92976;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RA "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054259; AAC15732.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR SMR; O92976; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_core.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_core.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR002531; HCV_NSI.
DR InterPro; IPR007094; RNA_pol_PSvir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_NSI; 1.
DR Pfam; PF01560; HCV_NSI; 1.
DR Pfam; PF01538; HCV_NSI; 1.
DR Pfam; PF02907; HCV_NSI; 1.
DR Pfam; PF01006; HCV_NSI; 1.
DR Pfam; PF01001; HCV_NSI; 1.
```

```

DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF01538; HCV_NS1; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW NON_TER
FT
SQ
SEQUENCE 2864 AA; C6DBE415F4A1D1D6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 147
Q9WLK8_9HEPC
ID Q9WLK8_9HEPC PRELIMINARY; PRT; 2864 AA.
AC Q9WLK8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244:161-172(1998).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054255; AAC15729.1; -; Genomic_RNA.
DR FIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q02828; 1NB4.
DR SMR; Q9WLK8; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008036; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.

```

```

DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW NON_TER
FT
SQ
SEQUENCE 2864 AA; C8AF2D0D7AE597E5 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 148
Q9WLK9_9HEPC
ID Q9WLK9_9HEPC PRELIMINARY; PRT; 2864 AA.
AC Q9WLK9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
infectious in vivo.";
RL Virology 244:161-172(1998).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).

```

```

RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93224885; PubMed=8385694;
RX Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RL diversity."
RL J. Gen. Virol. 74:661-668 (1993).
DR EMBL; AF054254; AAC15728.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR SMR; Q9WLK9; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008266; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F: serine-type peptidase activity; IEA.
DR GO; GO:0008236; F: structural molecule activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; P: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S23.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW NON_TER 2864 2864
FT SEQUENCE 2864 AA; 310553 MW; 208C0F9F2940BEBA CRC64;
SQ SEQUENCE 2864 AA; 310553 MW; 208C0F9F2940BEBA CRC64;

Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 149
Q9WLL0_9HEPC
ID Q9WLL0_9HEPC PRELIMINARY; PRT; 2864 AA.
AC Q9WLL0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
```

```

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo."
RL Virology 244:161-172 (1998).
DR [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity."
RL J. Gen. Virol. 72:2805-2809 (1991).
DR EMBL; AF054252; AAC15726.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; INB4.
DR SMR; Q9WLL0; 1029-1657, 2008-2170, 2420-2864.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; P: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S23.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW NON_TER 2864 2864
FT SEQUENCE 2864 AA; 310560 MW; 53464C5C744191D9 CRC64;
SQ SEQUENCE 2864 AA; 310560 MW; 53464C5C744191D9 CRC64;
```

```
Query Match 10.2%; Score 12; DB 2; Length 2864;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCL 29
    |||||
Db 1662 GGVLAALAAAYCL 1675

RESULT 150
O9J3F4_9HEPC
ID O9J3F4_9HEPC PRELIMINARY; PRT; 3008 AA.
AC O9J3F4;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
GN Name=MD34;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD34;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=158209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojiwa T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF208024; AAF61205.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0329; PQ0329.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q8J3F4; 1027-1655, 2006-2168, 2418-2947.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_039 HCV NS2.
DR InterPro; IPR004109; Peptidase S22.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV env; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
```

```
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3008 AA; 326838 MW; 99AE09E14C3109F4 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3008;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCL 29
    |||||
Db 1662 GGVLAALAAAYCL 1673

RESULT 151
POLG_HCVBK
ID POLG_HCVBK STANDARD; PRT; 3009 AA.
AC P26663;
DT 01-AUG-1992 (Rel. 23, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C)
DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
DE E2 (NS1) (gp58) (gp70); p7; Protease NS3-3 (EC 3.4.22.-) (p23); Serine
DE protease/NTpase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.1.-)
DE (Hepacivirin) (NS3P) (p70); Nonstructural protein 4A (NS4A) (p8);
DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
OS Hepatitis C virus (isolate BK) (HCV)
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111105;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=91140698; PubMed=1847440;
RA Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,
RA Onishi E., Andoh T., Yoshida I., Okayama H.;
RT "Structure and organization of the hepatitis C virus genome isolated
RT from human carriers.";
RL J. Virol. 65:1105-1113(1991).
RN [2]
RP PROTEIN SEQUENCE OF 1486-1499.
RX MEDLINE=96235224; PubMed=8647104;
RA Borowski P., Heiland M., Oehlmann K., Becker B., Kornetky L.;
RA "Non-structural protein 3 of hepatitis C virus inhibits
RT phosphorylation mediated by cAMP-dependent protein kinase.";
RL Eur. J. Biochem. 237:611-618(1996).
RN [3]
RP SUBCELLULAR LOCATION, AND RNA BINDING ACTIVITY OF CORE PROTEIN.
RX PubMed=8189501;
RA Santolini E., Migliaccio G., La Monica N.;
RA "Biosynthesis and biochemical properties of the hepatitis C virus core
RT protein.";
RL J. Virol. 68:3631-3641(1994).
RN [4]
RP CHARACTERIZATION OF PROTEASE NS2-3.
RX PubMed=9261354;
RA Pieroni L., Santolini E., Pipaldini C., Pacini L., Migliaccio G.,
RA La Monica N.;
RT "In vitro study of the NS2-3 protease of hepatitis C virus.";
RL J. Virol. 71:6373-6380(1997).
RN [5]
RP FUNCTION OF NS5A.
RX PubMed=9710605;
RA Gale M.J. Jr., Blakely C.M., Kwieciszewski B., Tan S.-L., Dossett M.,
RA Tang N.M., Korth M.J., Polyak S.J., Gretch D.R., Katze M.G.;
RT "Control of PKR protein kinase by hepatitis C virus nonstructural 5A
RT protein: molecular mechanisms of kinase regulation.";
```

- RL Mol. Cell. Biol. 18:5208-5218 (1998).
RN [6]
RP INTERACTION OF NS5A WITH HUMAN GRB2, AND MUTAGENESIS OF PRO-2321;
RX PRO-2322 AND PRO-2325.
RX MEDLINE=99254075; PubMed=10318919; DOI=10.1073/pnas.96.10.5533;
RA Tan S.-L., Nakao H., He Y., Vijayseri S., Neddermann P., Jacobs B.L.,
RA Mayer B.J., Katze M.G.;
RT "NS5A, a nonstructural protein of hepatitis C virus, binds growth
RT factor receptor-bound protein 2 adaptor protein in a Src homology 3
RT domain/ligand-dependent manner and perturbs mitogenic signaling.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:5533-5538 (1999).
RN [7]
RN MUTAGENESIS OF SER-2193.
RP PubMed=11118372; DOI=10.1006/viro.2000.0662;
RX Katze M.G., Kwiciszewski B., Goodlett D.R., Blakely C.M.,
RA Neddermann P., Tan S.-L., Abersold R.;
RA "Ser(2194) is a highly conserved major phosphorylation site of the
RT hepatitis C virus nonstructural protein NS5A";
RL Virology 278:501-513 (2000).
RN [8]
RX CHARACTERIZATION OF PROTEASE NS2-3.
RP MEDLINE=2158263; PubMed=11591719; DOI=10.1074/jbc.M108266200;
RA Thibault D., Maurice R., Pilote L., Lamarre D., Pause A.;
RT "In vitro characterization of a purified NS2/3 protease variant of
RT hepatitis C virus.";
RL J. Biol. Chem. 276:46678-46684 (2001).
RN [9]
RP INTERACTION OF NS5A WITH HUMAN PIK3R1.
RX PubMed=12186904; DOI=10.1128/JVI.76.18.9207-9217.2002;
RA He Y., Nakao H., Tan S.-L., Polyak S.J., Neddermann P., Vijayseri S.,
RA Jacobs B.L., Katze M.G.;
RT "Subversion of cell signaling pathways by hepatitis C virus
RT nonstructural 5A protein via interaction with Grb2 and FRS
RT phosphatidylinositol 3-kinase.";
RL J. Virol. 76:9207-9217 (2002).
RN [10]
RX CD81-BINDING AND HVR2 REGIONS.
RP MEDLINE=22547000; PubMed=12680945; DOI=10.1086/3682221;
RA Hofmann W.P., Sarrazin C., Kronenberger B., Schonberger B., Bruch K.,
RA Zeuzem S.;
RT "Mutations within the CD81-binding sites and hypervariable region 2 of
RT the envelope 2 protein: correlation with treatment response in
RT hepatitis C virus-infected patients.";
RL J. Infect. Dis. 187:982-987 (2003).
RN [11]
RX PHOSPHORYLATION OF NS5A.
RP PubMed=15016873; DOI=10.1128/JVI.78.7.3502-3513.2004;
RA Coito C., Diamond D.L., Neddermann P., Korth M.J., Katze M.G.;
RT "High-throughput screening of the yeast kinome: identification of
RT human serine/threonine protein kinases that phosphorylate the
RT hepatitis C virus NS5A protein.";
RL J. Virol. 78:3502-3513 (2004).
RN [12]
RX ISDR AND V3 REGIONS.
RP PubMed=15258967; DOI=10.1002/jmv.20144;
RA Vuilleumoz I., Khattab E., Sablon E., Ottevaere I., Durantel D.,
RA Vieux C., Trepo C., Zoulim F.;
RT "Genetic variability of hepatitis C virus in chronically infected
RT patients with viral breakthrough during interferon-ribavirin
RT therapy.";
RL J. Med. Virol. 74:41-53 (2004).
RN [13]
RX INTERACTION OF NS5A WITH CELLULAR PROTEINS.
RP PubMed=15607035;
RA Ahn J., Chung K.S., Kim D.U., Won M., Kim K.S., Nam M.,
RA Choi S.J., Kim H., Yoon M., Chae S.K., Hoe K.L.;
RT "Systematic identification of hepatocellular proteins interacting with
RT NS5A of the hepatitis C virus.";
RL J. Biochem. Mol. Biol. 37:741-748 (2004).
RN [14]
RX REVIEW.
RP PubMed=10718937; DOI=10.1046/j.1365-2893.2000.00201.x;
RA McLauchlan J.;
RT "Properties of the hepatitis C virus core protein: a structural
RT protein that modulates cellular processes.";
RL J. Viral Hepat. 7:2-14 (2000).
RN [15]
RP REVIEW.
RX PubMed=14752815; DOI=10.1002/hep.20032;
RA Penin F., Dublsson J., Rey F.A., Moradpour D., Pawlotsky J.-M.;
RT "Structural biology of hepatitis C virus.";
RL Hepatology 39:5-19 (2004).
RN [16]
RX X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1026-1214.
RP MEDLINE=97015088; PubMed=8861916; DOI=10.1016/S0092-8674(00)81350-1;
RA Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
RA Moomaw E.W., Adachi T., Hostomska Z.;
RT "The crystal structure of hepatitis C virus NS3 proteinase reveals a
RT trypsin-like fold and a structural zinc binding site.";
RL Cell 87:331-342 (1996).
RN [17]
RX X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1026-1205 AND 1677-1690.
RP MEDLINE=98227846; PubMed=9568891;
RA Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
RA Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
RT "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
RT virus: a 2.2-A resolution structure in a hexagonal crystal form.";
RL Protein Sci. 7:837-847 (1998).
RN [18]
RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 1215-1649.
RP MEDLINE=98279021; PubMed=9614113; DOI=10.1074/jbc.273.24.15045;
RA Cho H.-S., Ha N.-C., Kang L.-W., Chung K.M., Back S.H., Jang S.K.,
RA Oh B.-H.;
RT "Crystal structure of RNA helicase from genotype 1b hepatitis C virus.
RT A feasible mechanism of unwinding duplex RNA.";
RL J. Biol. Chem. 273:15045-15052 (1998).
RN [19]
RX X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 1012-1656.
RP MEDLINE=20045185; PubMed=10574797; DOI=10.1016/S0969-2126(00)80025-8;
RA Yao N., Reichert P., Taremi S.S., Prosser W.W., Weber P.C.;
RT "Molecular views of viral polypeptide processing revealed by the
RT crystal structure of the hepatitis C virus bifunctional protease-
RT helicase.";
RL Structure 7:1353-1363 (1999).
RN [20]
RX STRUCTURE BY NMR OF 1026-1205.
RP PubMed=10366511; DOI=10.1006/jmbi.1999.2745;
RA Barbato G., Cicero D.O., Nardi M.C., Steinkuehler C., Cortese R.,
RA De Francesco R., Bazzo R.;
RT "The solution structure of the N-terminal proteinase domain of the
RT hepatitis C virus (HCV) NS3 protein provides new insights into its
RT activation and catalytic mechanism.";
RL J. Mol. Biol. 289:371-384 (1999).
RN [21]
RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 2419-2949.
RP MEDLINE=20027495; PubMed=10557268; DOI=10.1073/pnas.96.23.13034;
RA Bressanelli S., Tomei L., Rousset A., Incitti I., Vitale R.L.,
RA Mathieu M., De Francesco R., Rey F.A.;
RT "Crystal structure of the RNA-dependent RNA polymerase of hepatitis C
RT virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13034-13039 (1999).
RN [22]
RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2413-2988.
RP PubMed=10504728; DOI=10.1038/13305;
RA Leeburg C.A., Cable M.B., Ferrari E., Hong Z., Mannarino A.F.,
RA Weber P.C.;
RT "Crystal structure of the RNA-dependent RNA polymerase from hepatitis
RT C virus reveals a fully encircled active site.";
RL Nat. Struct. Biol. 6:937-943 (1999).
RN [23]
RX X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 2419-2998.
RP MEDLINE=20045191; PubMed=10574802; DOI=10.1016/S0969-2126(00)80031-3;
RA Ago H., Adachi T., Yoshida A., Yamamoto M., Habuka N., Yatsunami K.,
RA Miyano M.;
RT "Crystal structure of the RNA-dependent RNA polymerase of hepatitis C
RT virus.";

RL Structure 7:1417-1426(1999).
RN [24]
Query Match 10.2%; Score 12; DB 1; Length 3009;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1663 GGVLAALAAAYCL 1674
RESULT 152
POLG_HCVCO
ID POLG_HCVCO STANDARD; PRT; 3009 AA.
AC Q9WMX2;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Genome polyprotein (Contains: Core protein p19 (Capsid protein C)
DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
DE E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine
DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.-)
DE (Hepacivirin) (NS3p) (p70); Nonstructural protein 4A (NS4A) (p8);
DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
OS Hepatitis C virus (isolate Con1) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=333284;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=99322193; PubMed=10390360; DOI=10.1126/science.285.5424.110;
RA Lohmann V., Koerner F., Koch J.O., Herian U., Theilmann L.,
RA Bartenschlager R.;
RT "Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
RT line.";
RL Science 285:110-113(1999).
RN [2]
RP FUNCTION OF NS3 AND NS4A.
RX PubMed=12702807; DOI=10.1126/science.1082604;
RA Foy E., Li K., Wang C., Sumpter R. Jr., Ikeda M., Lemon S.M.,
RA Gale M. Jr.;
RT "Regulation of interferon regulatory factor-3 by the hepatitis C virus
RT serine protease.";
RL Science 300:1145-1148(2003).
RN [3]
RP FUNCTION OF NS5A.
RX PubMed=15542681; DOI=10.1126/JVI.78.23.13306-13314.2004;
RA Neddermann P., Quintavalle M., Di Pietro C., Clementi A.,
RA Cerretti M., Altamura S., Bartholomew L., De Francesco R.;
RT "Reduction of hepatitis C virus NS5A hyperphosphorylation by selective
RT inhibition of cellular kinases activates viral RNA replication in cell
RT culture.";
RL J. Virol. 78:13306-13314(2004).
RN [4]
RP MUTAGENESIS OF SER-2157; SER-2172; SER-2178; SER-2193; SER-2196;
RP SER-2199; SER-2200; SER-2201; SER-2203; SER-2206; SER-2209; SER-2245;
RP SER-2254 AND SER-2268.
RX PubMed=15709040; DOI=10.1126/JVI.79.5.3187-3194.2005;
RA Appel N., Pietschmann T., Bartenschlager R.;
RT "Mutational analysis of hepatitis C virus nonstructural protein 5A:
RT potential role of differential phosphorylation in RNA replication and
RT identification of a genetically flexible domain.";
RL J. Virol. 79:3187-3194(2005).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 2007-2169, AND DISULFID BOND.
RX PubMed=15902263; DOI=10.1038/nature03580;
RA Tellinghuisen T.L., Marcotrigiano J., Rice C.M.;
RT "Structure of the zinc-binding domain of an essential component of the
RT hepatitis C virus replicase.";
RL Nature 435:374-379(2005).
CC -1- FUNCTION: Core protein packages viral RNA to form a viral

nucleocapsid, and promotes virion budding. Modulates viral translation initiation by interacting with HCV IRES and 40S ribosomal subunit. Also regulates many host cellular functions such as signaling pathways and apoptosis. Prevents the establishment of cellular antiviral state by blocking the interferon-alpha/beta (IFN-alpha/beta) and IFN-gamma signaling pathways and by inducing human STAT1 degradation. Plays an important role in virus-mediated cell transformation leading to hepatocellular carcinomas. Interacts with, and activates STAT3 leading to cellular transformation. May repress the promoter of p53, and sequester CREB3 and Sp110 isoform 3/Sp110b in the cytoplasm. Also represses cell cycle negative regulating factor CDKN1A, thereby interrupting an important checkpoint of normal cell cycle regulation. Targets transcription factors involved in the regulation of inflammatory responses and in the immune response; suppresses NK-kappaB activation, and activates AP-1. Mediates apoptotic pathways through association with TNF-type receptors TNFRSF1A and LTBR, although its effect on death receptor-induced apoptosis remains controversial. Enhances TRAIL mediated apoptosis, suggesting that it might play a role in immune-mediated liver cell injury. Secreted core protein is able to bind ClO81 at the T-cell surface, resulting in down-regulation of T-lymphocytes proliferation. May transactivate human MYC, Rous sarcoma virus LTR, and SV40 promoters. May suppress the human FOS and HIV-1 LTR activity. May alter lipid metabolism by interacting with hepatocellular proteins involved in lipid accumulation and storage (By similarity).
-1- FUNCTION: Envelope glycoproteins E1 and E2 are involved in virus attachment to the host cell as well as in virus endocytosis and fusion with host membrane. E2 inhibits human PKR activation, preventing the establishment of an antiviral state (By similarity).
-1- FUNCTION: p7 seems to be a hexameric ion channel protein (viroporin) and is inhibited by the antiviral drug amantadine. Also inhibited by long-alkyl-chain iminosugar derivatives.
-1- FUNCTION: Protease NS2-3, which is a putative cysteine protease, is responsible for the autocatalytic cleavage of NS2-NS3. Seems to undergo self-inactivation following maturation (By similarity).
-1- FUNCTION: NS3 displays three distinct enzymatic activities: serine protease, NTPase and RNA helicase. NS3 serine protease, in association with NS4A, is responsible for the cleavages of NS3-NS4A, NS4A-NS4B, NS4B-NS5A and NS5A-NS5B. NS3 RNA helicase binds to RNA and unwinds dsRNA in the 3' to 5' direction (By similarity). NS3/NS4A complex also prevents phosphorylation of human IRE3, thus preventing the establishment of dsRNA induced antiviral state.
-1- FUNCTION: NS4B may induce a specific membrane alteration that serves as a scaffold for the virus replication complex. This membrane alteration gives rise to the so called ER-derived membranous web that contains the replication complex (By similarity).
-1- FUNCTION: NS5A is a component of the replication complex. Downregulates viral IRES translation initiation. Mediates interferon resistance, presumably by interacting with and inhibiting human PKR/p38. The hyperphosphorylated form of NS5A is an inhibitor of viral replication (By similarity).
-1- FUNCTION: NS5B is a RNA-dependent RNA polymerase that plays an essential role in the virus replication (By similarity).
-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the p6 position, Cys or Thr in p1 and Ser or Ala in p1'.
-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
-1- COFACTOR: Binds 1 zinc ion per NS3 protease domain. Binds 1 zinc ion per NS5A N-terminal domain (By similarity).
-1- ENZYME REGULATION: Activity of protease NS2-3 is dependent on zinc ions and completely inhibited by EDTA. Serine protease NS3 is also activated by zinc ions (By similarity).
-1- SUBUNIT: Core protein is a homodimer that binds the C-terminal part of E1. Interacts with numerous cellular proteins. Interacts with human STAT1, inducing its degradation and human STAT3, constitutively activating it. Associates with human LTBR and

TNFRSF1A receptors and possibly induces apoptosis. Binds to human SP110 isoform 3/Sp110b, HNRPK, CLOR1, YWHAZ, DDX3X, APOA2 and RXRA proteins. Interacts with human CREB3, nuclear transcription protein, triggering cell transformation. May interact with human p53. Also binds human cytochromes KRT8, KRT18, KRT19 and VIM (vimentin). E1 and E2 glycoproteins form a heterodimer that binds to human LDLR, CD81 and SCAR1 receptors, but this binding is not sufficient for infection, some additional liver specific cofactors may be needed. E2 binds and inhibits human PRKR. p7 forms a homohexamer. NS2 forms a homodimer and interacts with all other nonstructural (NS) proteins. NS4A interacts with NS3 serine protease and stabilizes it. NS3-NS4A complex is essential for the activation of the latter and allows membrane anchorage of NS3. NS4B forms a homodimer and interacts with all other NS proteins. NS5A forms a homodimer and interacts with all other nonstructural (NS) proteins. NS5A also interacts with human PRKR, GRB2, PIK3R1 and with most Src-family kinases. NS5B is a homooligomer (By similarity).

-I- SUBCELLULAR LOCATION: The virion assembly and budding occurs from the ER membrane. The N-terminal transmembrane domains of the polyprotein probably possess an endoplasmic reticulum (ER) signal sequence function in their C-termini, leading the nascent polyprotein to the ER membrane. After cleavage by host signal peptidase, these ER signal peptides retain at the C-terminus of the concerned proteins (core, E1, E2 and p7), serving as ER membrane anchors. Core protein is cytoplasmic. It is also located on mitochondrial and endoplasmic reticulum membranes and at the surface of lipid droplets. A minor proportion is present in the nucleus. An unknown proportion is secreted. E1, E2, NS2 and NS4B are integral ER membrane proteins. The C-terminal transmembrane domains of envelope glycoproteins E1 and E2 form a hairpin structure before cleavage by host signal peptidase. A reorientation of the second hydrophobic stretch occurs after cleavage producing a single reoriented transmembrane domain. These events explain the final topology of these proteins. ER retention of E1 and E2 is leaky and, in overexpression conditions, a small fraction of both proteins reaches the plasma membrane. NS3 is associated to the ER membrane through its binding to NS4A. NS5A Membrane insertion of the membrane-anchored proteins NS4A, NS5A and NS5B occurs after processing by the NS3 protease. NS5A is perinuclear. A fraction of p7 localizes to the plasma membrane (By similarity).

-I- DOMAIN: The transmembrane regions of envelope E1 and E2 glycoproteins are involved in heterodimer formation. ER localization, and assembly of these proteins. Envelope E2 glycoprotein contain two highly variable regions called hypervariable region 1 and 2 (HVR1 and HVR2) and two CD81-binding sites. HVR1 is implicated in the SCAR1-mediated cell entry. HVR2 and CD81-binding sites may be involved in sensitivity and/or resistance to IFN-alpha therapy (By similarity).

-I- DOMAIN: The C-terminus of NS5A contains two variable regions called interferon sensitivity determining region (ISDR) and variable region 3 (V3). ISDR and V3 may be involved in sensitivity and/or resistance to IFN-alpha therapy (By similarity).

-I- DOMAIN: The N-terminal one-third of serine protease NS3 contains the protease activity. This region contains a zinc atom that does not belong to the active site, but may play a structural rather than a catalytic role. This region is essential for the activity of protease NS2-3, maybe by contributing to the folding of the latter. The helicase activity is located in the C-terminus (By similarity).

-I- PM: Specific enzymatic cleavages in vivo yield mature proteins. The structural proteins, core, E1, E2 and p7 are produced by proteolytic processing by host signal peptidases. The core protein is synthesized as a 21 kDa precursor which is retained in the ER membrane through a hydrophobic propeptide, which is cleaved to release the 19 kDa mature core protein. The other proteins are cleaved by the viral proteases (By similarity).

Query Match 10.2%; Score 12; DB 1; Length 3009;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1663 GGVLAALAAAYCL 1674
RESULT 153
POLG HCVAJ STANDARD; PRT; 3009 AA.
ID P26662; P89966; Q81755;
DT 01-AUG-1992 (Rel. 23, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C) (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.1.-) (Hepacivirin) (NS3p) (p70); Nonstructural protein 4A (NS4A) (p8); Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A) (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p58)].
OS Hepatitis C virus (isolate Japanese) (HCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
OC NCBI_TaxID=11116;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RX MEDLINE=9108550; PubMed=2175903;
RA Kato N., Hijikata M., Nakagawa Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
RT "Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis."
RL Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
RN [2]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=91192160; PubMed=1849488; DOI=10.1016/0014-5793(91)80322-T;
RA Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K., Ohkoshi S., Shimotohno K.;
RT "Molecular structure of the Japanese hepatitis C viral genome."
RL FEBS Lett. 280:325-328(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
RA Tanaka T.;
RN Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE [GENOMIC RNA] OF 722-1907, IDENTIFICATION OF BOTH PROTEASES, AND MUTAGENESIS OF CYS-921; HIS-931; HIS-951; GLU-971; GLU-979; CYS-992; GLU-1008; GLU-1057; HIS-1082; ASP-1106; CYS-1122; CYS-1124; HIS-1135; SER-1164; CYS-1170; HIS-1174; CYS-1184; GLU-1198; GLU-1201; HIS-1226 AND HIS-1228.
RX MEDLINE=93323208; PubMed=8392606;
RA Hijikata M., Mizushima H., Akagi T., Mori S., Kakiuchi N., Kato N., Tanaka T., Kimura K., Shimotohno K.;
RT "Two distinct protease activities required for the processing of a putative nonstructural precursor protein of hepatitis C virus."
RL J. Virol. 67:4665-4675(1993).
RN [5]
RP INHIBITION OF HEPATITIS B VIRUS GENE EXPRESSION.
RX PubMed=8396658;
RA Shih C.-M., Lo S.-J., Miyamura T., Chen S.-Y., Lee Y.-H.W.;
RT "Suppression of hepatitis B virus expression and replication by hepatitis C virus core protein in HuH-7 cells."
RL J. Virol. 67:5823-5832(1993).
RN [6]
RP PROCESSING OF POLYPROTEIN, MUTAGENESIS OF HIS-1082; ASP-1106 AND SER-1164, AND IDENTIFICATION OF THE SERINE PROTEASE.
RX PubMed=8389908;
RA Bartenschlager R., Ahlborn-Laake L., Mous J., Jacobsen H.;
RT "Nonstructural protein 3 of the hepatitis C virus encodes a serine-type proteinase required for cleavage at the NS3/4 and NS4/5 junctions."
RL J. Virol. 67:3835-3844(1993).
RN [7]
RP PHOSPHORYLATION OF NS5A.
RX PubMed=7999043;

RA Kaneko T., Tanji Y., Satoh S., Hijikata M., Asabe S., Kimura K.,
RA Shimotohno K.;
RT "Production of two phosphoproteins from the NS5A region of the
RT hepatitis C viral genome.";
RL Biochem. Biophys. Res. Commun. 205:320-326(1994).
RN [8]
RP MUTAGENESIS OF SER-2193; SER-2196; SER-2199; SER-2200; SER-2201;
RP SER-2203; SER-2206; SER-2209 AND SER-2220.
RX PubMed=7769656;
RA Tanji Y., Kaneko T., Satoh S., Shimotohno K.;
RT "Phosphorylation of hepatitis C virus-encoded nonstructural protein
RT NS5A.";
RL J. Virol. 69:3980-3986(1995).
RN [9]
RP FUNCTION OF NS5A.
RX MEDLINE=96127954; PubMed=8531962; DOI=10.1056/NEJM19960113340203;
RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
RA Yamamoto C., Ogura Y., Izumi N., Maruno P., Sato C.;
RT "Mutations in the nonstructural protein 5A gene and response to
RT interferon in patients with chronic hepatitis C virus 1b infection.";
RL N. Engl. J. Med. 334:77-81(1996).
RN [10]
RP FUNCTION OF NS5A.
RX PubMed=9343247;
RA Kato N., Imai K.H., Ono-Nita S.K., Shiratori Y., Omata M.;
RT "Hepatitis C virus nonstructural region 5A protein is a potent
RT transcriptional activator.";
RL J. Virol. 71:8856-8859(1997).
RN [11]
RP ZINC-BINDING REGION OF SERINE PROTEASE NS3, COPACTOR, AND MUTAGENESIS
RP OF CYS-1041; CYS-1072; CYS-1077; HIS-1082; HIS-1135; SER-1164;
RP HIS-1174; CYS-1184; HIS-1226 AND HIS-1228.
RX PubMed=9060645;
RA Stempiak M., Hostomska Z., Nodes B.R., Hostomsky Z.;
RT "The NS3 proteinase domain of hepatitis C virus is a zinc-containing
RT enzyme.";
RL J. Virol. 71:2881-2886(1997).
RN [12]
RP SUBCELLULAR LOCATION OF CORE PROTEIN.
RX PubMed=9621068;
RA Yasui K., Wakita T., Tsukiyama-Kohara K., Funahashi S.-I.,
RA Ichikawa M., Kajita T., Moradpour D., Wands J.R., Kohara M.;
RT "The native form and maturation process of hepatitis C virus core
RT protein.";
RL J. Virol. 72:6048-6055(1998).
RN [13]
RP ISDR REGION.
RX PubMed=9525599;
RA Pawlotsky J.-M., Germanidis G., Neumann A.U., Pellerin M.,
RA Frainais P.-O., Dhumeaux D.;
RT "Interferon resistance of hepatitis C virus genotype 1b: relationship
RT to nonstructural 5A gene quasi-species mutations.";
RL J. Virol. 72:2795-2805(1998).
RN [14]
RP OLIGOMERIZATION OF NS5B.
RX PubMed=11907226; DOI=10.1128/JVI.76.8.3865-3872.2002;
RA Wang Q.M., Hockman M.A., Staschke K., Johnson R.B., Case K.A., Lu J.,
RA Parsons S., Zhang F., Rathnachalam R., Kirkegaard K., Colacino J.M.;
RT "Oligomerization and cooperative RNA synthesis activity of hepatitis C
RT virus RNA-dependent RNA polymerase.";
RL J. Virol. 76:3865-3872(2002).
RN [15]
RP INHIBITION OF P7 BY AMANTADINE.
RX MEDLINE=22448611; PubMed=12560074; DOI=10.1016/S0014-5793(02)03851-6;
RA Griffin S.D., Beales L.P., Clarke D.S., Worsfold O., Evans S.D.,
RA Jaeger J., Harris M.P., Rowlands D.J.;
RT "The p7 protein of hepatitis C virus forms an ion channel that is
RT blocked by the antiviral drug, Amantadine.";
RL FEBS Lett. 535:34-38(2003).
RN [16]
RP INTERACTION OF CORE PROTEIN WITH HUMAN SP110.
RX PubMed=14559998; DOI=10.1128/MCB.23.21.7498-7509.2003;
RA Wataashi K., Hijikata M., Tagawa A., Doi T., Marusawa H.,

RA Shimotohno K.;
RT "Modulation of retinoid signaling by a cytoplasmic viral protein via
RT sequestration of Sp10b, a potent transcriptional corepressor of
RT retinoic acid receptor, from the nucleus.";
RL Mol. Cell. Biol. 23:7498-7509(2003).
RN [17]
RP CLEAVAGE OF CORE PROTEIN BY THE SIGNAL PEPTIDE PEPTIDASE, AND
RP MUTAGENESIS OF LEU-138; VAL-139; LEU-143; ILE-175; PHE-176; LEU-177;
RP LEU-178; LEU-180; SER-182 AND CYS-183.
RX PubMed=15163730; DOI=10.1128/JVI.78.12.6370-6380.2004;
RA Okamoto K., Moriishi K., Miyamura T., Matsuura Y.;
RT "Intramembrane proteolysis and endoplasmic reticulum retention of
RT hepatitis C virus core protein.";
RL J. Virol. 78:6370-6380(2004).
RN [18]
RP FUNCTION OF CORE PROTEIN.
RX PubMed=15760888; DOI=10.1074/jbc.M501826200;
RA Boni S., Laverigne J.-P., Boulant S., Cahour A.;
RT "Hepatitis C virus core protein acts as a trans-modulating factor on
RT internal translation initiation of the viral RNA.";
RL J. Biol. Chem. 280:17737-17748(2005).
RN [19]
RP INTERACTION OF CORE PROTEIN WITH HUMAN STAT1.
RX PubMed=15825084;
RA Lin W., Choe W.H., Hiasa Y., Kamegaya Y., Blackard J.T., Schmidt E.V.,
RA Chung R.T.;
RT "Hepatitis C virus expression suppresses interferon signaling by
RT degrading STAT1.";
RL Gastroenterology 128:1034-1041(2005).
RN [20]
RP SUBCELLULAR LOCATION OF CORE PROTEIN.
RX PubMed=15613354; DOI=10.1128/JVI.79.2.1271-1281.2005;
RA Suzuki R., Sakamoto S., Teutsami T., Rikimaru A., Tanaka K.,
RA Shimoiike T., Moriishi K., Iwasaki T., Mizumoto K., Matsuura Y.,
RA Miyamura T., Suzuki T.;
RT "Molecular determinants for subcellular localization of hepatitis C
RT virus core protein.";
RL J. Virol. 79:1271-1281(2005).
RN [21]
RP REVIEW.
RX PubMed=10718937; DOI=10.1046/j.1365-2893.2000.00201.x;
RA McLauchlan J.;
RT "Properties of the hepatitis C virus core protein: a structural
RT protein that modulates cellular processes.";
RL J. Viral Hepat. 7:2-14(2000).
RN [22]
RP REVIEW.
RX PubMed=14752815; DOI=10.1002/hep.20032;
RA Penin F., Dubuisson J., Rey F.A., Moradpour D., Pawlotsky J.-M.;
RT "Structural biology of hepatitis C virus.";
RL Hepatology 39:5-19(2004).
RN [23]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1026-1212.
RX MEDLINE=20167193; PubMed=10702283; DOI=10.1074/jbc.275.10.7152;
RA Di Marco S., Rizzi M., Volpari C., Waleh M.A., Narjes F.,
RA Collaruso S., De Francesco R., Matassa V.G., Sollazzo M.;
RT "Inhibition of the hepatitis C virus NS3/4A protease. The crystal
RT structures of two protease-inhibitor complexes.";
RL J. Biol. Chem. 275:7152-7157(2000).
RN [24]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 1026-1212 IN COMPLEX WITH A

Query Match 10.24; Score 12; DB 1; Length 3009;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1663 GGVLAALAAAYCL 1674

RESULT 154
POLG_HCVJT

ID POLG HCVUT STANDARD; PRT; 3009 AA.
 AC Q00269;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C)
 DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
 DE E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine
 DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.-)
 DE (Hepadacivirin) (NS3p) (p70); Nonstructural protein 4A (NS4A) (p8);
 DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
 DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
 OS Hepatitis C virus (isolate HC-JT) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OC NCBI_TaxID=31642;
 RX MEDLINE=92295714; PubMed=1318627; DOI=10.1016/0168-1702(92)90066-1;
 RA Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T.,
 RA Hijikata M., Ishimura Y., Shimotohno K.;
 RT "Molecular cloning of hepatitis C virus genome from a single Japanese
 RT carrier: sequence variation within the same individual and among
 RT infected individuals.";
 RL Virus Res. 23:39-53(1992).
 CC -1- FUNCTION: Core protein packages viral RNA to form a viral
 CC nucleocapsid, and promotes virion budding. Modulates viral
 CC translation initiation by interacting with HCV IRES and 40S
 CC ribosomal subunit. Also regulates many host cellular functions
 CC such as signaling pathways and apoptosis. Prevents the
 CC establishment of cellular antiviral state by blocking the
 CC interferon-alpha/beta (IFN-alpha/beta) and IFN-gamma signaling
 CC pathways and by inducing human STAT1 degradation. Plays an
 CC important role in virus-mediated cell transformation leading to
 CC hepatocellular carcinomas. Interacts with, and activates SPAT3
 CC leading to cellular transformation. May repress the promoter of
 CC p53, and sequester CREB3 and Sp110 isoform 3/Sp110b in the
 CC cytoplasm. Also represses cell cycle negative regulating factor
 CC CDKN1A, thereby interrupting an important check point of normal
 CC cell cycle regulation. Targets transcription factors involved in
 CC the regulation of inflammatory responses and in the immune
 CC response; suppresses NK-kappaB activation, and activates AP-1.
 CC Mediates apoptotic pathways through association with TNF-type
 CC receptors TNFRSF1A and LTR, although its effect on death
 CC receptors-induced apoptosis remains controversial. Enhances TRAIL
 CC mediated apoptosis, suggesting that it might play a role in
 CC immune-mediated liver cell injury. Secreted core protein is able
 CC to bind C10R1 at the T-cell surface, resulting in down-regulation
 CC of T-lymphocytes proliferation. May transactivate human MYC, Rous
 CC sarcoma virus LTR, and SV40 promoters. May suppress the human FOS
 CC and HIV-1 LTR activity. May alter lipid metabolism by interacting
 CC with hepatocellular proteins involved in lipid accumulation and
 CC storage (By similarity).
 CC -1- FUNCTION: Envelope glycoproteins E1 and E2 are involved in virus
 CC attachment to the host cell as well as in virus endocytosis and
 CC fusion with host membrane. E2 inhibits human PRKR activation,
 CC preventing the establishment of an antiviral state (By
 CC similarity).
 CC -1- FUNCTION: p7 seems to be a hexameric ion channel protein
 CC (viroporin) and is inhibited by the antiviral drug amantadine.
 CC Also inhibited by long-alkyl-chain iminosugar derivatives.
 CC Essential for infectivity (By similarity).
 CC -1- FUNCTION: Protease NS2-3, which is a putative cysteine protease,
 CC is responsible for the autocatalytic cleavage of NS2-NS3. Seems to
 CC undergo self-inactivation following maturation (By similarity).
 CC -1- FUNCTION: NS3 displays three distinct enzymatic activities: serine
 CC protease, NTPase and RNA helicase. NS3 serine protease, in
 CC association with NS4A, is responsible for the cleavages of NS3-
 CC NS4A, NS4A-NS4B, NS4B-NS5A and NS5A-NS5B. NS3 RNA helicase binds
 CC to RNA and unwinds dsRNA in the 3' to 5' direction. NS3/NS4A
 CC complex also prevents phosphorylation of human IRF3, thus
 CC preventing the establishment of dsRNA induced antiviral state (By
 CC similarity).
 CC -1- FUNCTION: NS4B may induce a specific membrane alteration that
 CC serves as a scaffold for the virus replication complex. This
 CC membrane alteration gives rise to the so called ER-derived
 CC membranous web that contains the replication complex (By
 CC similarity).
 CC -1- FUNCTION: NS5A is a component of the replication complex.
 CC Downregulates viral IRES translation initiation. Mediates
 CC interferon resistance, presumably by interacting with and
 CC inhibiting human PRKR/PKR. The hyperphosphorylated form of NS5A is
 CC an inhibitor of viral replication (By similarity).
 CC -1- FUNCTION: NS5B is a RNA-dependent RNA polymerase that plays an
 CC essential role in the virus replication (By similarity).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the p6
 CC position, Cys or Thr in p1 and Ser or Ala in p1'.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -1- COFACTOR: Binds 1 zinc ion per NS3 protease domain. Binds 1 zinc
 CC ion per NS5A N-terminal domain (By similarity).
 CC -1- ENZYME REGULATION: Activity of protease NS2-3 is dependent on zinc
 CC ions and completely inhibited by EDTA. Serine protease NS3 is also
 CC activated by zinc ions (By similarity).
 CC -1- SUBUNIT: Core protein is a homomultimer that binds the C-terminal
 CC part of E1. Interacts with numerous cellular proteins. Interacts
 CC with human STAT1, inducing its degradation and human SPAT3,
 CC constitutively activating it. Associates with human LTR and
 CC TNFRSF1A receptors and possibly induces apoptosis. Binds to human
 CC Sp110 isoform 3/Sp110b, HNRPK, C10R1, YWHA, DDX3X, APOA2 and RXRA
 CC proteins. Interacts with human CREB3 nuclear transcription
 CC protein, triggering cell transformation. May interact with human
 CC p53. Also binds human cytochromes KRT8, KRT18, KRT19 and VIM
 CC (vimentin). E1 and E2 glycoproteins form a heterodimer that binds
 CC to human LDLR, CD81 and SCARB1 receptors, but this binding is not
 CC sufficient for infection, some additional liver specific cofactors
 CC may be needed. E2 binds and inhibits human PRKR. p7 forms a
 CC homohexamer. NS2 forms a homodimer and interacts with all other
 CC nonstructural (NS) proteins. NS4A interacts with NS3 serine
 CC protease and stabilizes it. NS3-NS4A complex is essential for the
 CC activation of the latter and allows membrane anchorage of NS3.
 CC NS4B forms a homodimer and interacts with all other NS proteins.
 CC NS5A forms a homodimer and interacts with all other nonstructural
 CC (NS) proteins. NS5A also interacts with human PRKR, GRB2, PIK3R1
 CC and with most Src-family kinases. NS5B is a homooligomer (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: The virion assembly and budding occurs from
 CC the ER membrane. The N-terminal transmembrane domains of the
 CC polyprotein probably possess an endoplasmic reticulum (ER) signal
 CC sequence function in their C-termini, leading the nascent
 CC polyprotein to the ER membrane. After cleavage by host signal
 CC peptidase, these ER signal peptides retain at the C-terminus of
 CC the concerned proteins (core, E1, E2 and p7), serving as ER
 CC membrane anchors. Core protein is cytoplasmic. It is also located
 CC on mitochondrial and endoplasmic reticulum membranes and at the
 CC surface of lipid droplets. A minor proportion is present in the
 CC nucleus. An unknown proportion is secreted. E1, E2, NS2 and NS4B
 CC are integral ER membrane proteins. The C-terminal transmembrane
 CC domains of envelope glycoproteins E1 and E2 form a hairpin
 CC structure before cleavage by host signal peptidase. A
 CC reorientation of the second hydrophobic stretch occurs after
 CC cleavage producing a single reoriented transmembrane domain. These
 CC events explain the final topology of these proteins. ER retention
 CC of E1 and E2 is leaky and, in overexpression conditions, a small
 CC fraction of both proteins reaches the plasma membrane. NS3 is
 CC associated to the ER membrane through its binding to NS4A.
 CC Membrane insertion of the membrane-anchored proteins NS4A, NS5A
 CC and NS5B occurs after processing by the NS3 protease. NS5A is
 CC perinuclear. A fraction of p7 localizes to the plasma membrane (By
 CC similarity).
 CC -1- DOMAIN: The transmembrane regions of envelope E1 and E2
 CC glycoproteins are involved in heterodimer formation, ER
 CC localization, and assembly of these proteins. Envelope E2
 CC glycoprotein contain two highly variable regions called
 CC hypervariable region 1 and 2 (HVR1 and HVR2) and two CD81-binding

CC sites. HVR1 is implicated in the SCARBI-mediated cell entry. HVR2
 CC and CD81-binding sites may be involved in sensitivity and/or
 CC resistance to IFN-alpha therapy (By similarity).
 CC -1- DOMAIN: The C-terminus of NS5A contains two variable regions
 CC called interferon sensitivity determining region (ISDR) and
 CC variable region 3 (V3). ISDR and V3 may be involved in sensitivity
 CC and/or resistance to IFN-alpha therapy (By similarity).
 CC -1- DOMAIN: The N-terminal one-third of serine protease NS3 contains
 CC the protease activity. This region contains a zinc atom that does
 CC not belong to the active site, but may play a structural rather
 CC than a catalytic role. This region is essential for the activity
 CC of protease NS2-3, maybe by contributing to the folding of the
 CC latter. The helicase activity is located in the C-terminus (By
 CC similarity).
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC The structural proteins, core, E1, E2 and p7 are produced by
 CC proteolytic processing by host signal peptidases. The core protein
 CC is synthesized as a 21 kDa precursor which is retained in the ER
 CC membrane through a hydrophobic propeptide, which is cleaved to
 CC release the 19 kDa mature core protein. The other proteins are
 CC cleaved by the viral proteases (By similarity).
 CC -1- PTM: Envelope E1 and E2 glycoproteins are highly N-glycosylated
 CC (By similarity).
 CC -1- PTM: Core protein is phosphorylated by host PKC and PKA (By
 CC similarity).
 CC -1- PTM: NS5A is phosphorylated in a basal form termed p56. p58 is an
 CC hyperphosphorylated form of p56. p56 and p58 coexist in the cell
 CC in roughly equivalent amounts. Hyperphosphorylation is dependent
 CC on the presence of NS4B in cis. Human AKT1, RPS6KB1/p70S6K,
 CC MAP2K1/MEK1 and MAP2K6/MEK6 kinases may be responsible for NS5A
 CC phosphorylation (By similarity).
 CC -1- MISCELLANEOUS: Cell culture adaptation of the virus leads to
 CC mutations in NS5A, reducing its inhibitory effect on replication
 CC (By similarity).
 CC -1- MISCELLANEOUS: The virion is a nucleocapsid covered by a
 CC lipoprotein envelope. The nucleocapsid is composed of the core
 CC protein, forming an internal icosahedral coat that encapsidates
 CC the genomic RNA. The envelope contains two proteins: the envelope
 CC glycoproteins E1 and E2.
 CC -1- MISCELLANEOUS: Core protein exerts viral interference on hepatitis
 CC B virus when HCV and HBV coinfect the same cell, by suppressing
 CC HBV gene expression, RNA encapsidation and budding (By
 CC similarity).
 CC -1- SIMILARITY: Contains 1 peptidase C18 domain.
 CC -1- SIMILARITY: Contains 1 peptidase S29 domain.
 CC -1- CAUTION: There is a doubt concerning the orientation of the N-
 CC terminus of NS4B, which could be luminal with a 5th transmembrane
 CC segment. The C-terminus of NS2 may be luminal with a fourth
 CC transmembrane segment.
 CC -1- CAUTION: The core gene probably also codes for alternative reading
 CC frame proteins (ARFPs). Many functions depicted for the core
 CC protein might belong to the ARFPs.
 CC -1- DATABASE: NAME=suHCVdb; NOTE=The European HCV database";
 CC WWW="http://euHCVdb.ibcp.fr".

Query Match 10.2%; Score 12; DB 1; Length 3009;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 1663 GGVLAALAAAYCL 1674

RESULT 155
 POLG HCVTV STANDARD; PRT; 3009 AA.
 AC P29846;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C)
 DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein

DE E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.21.98) (p23); Serine
 DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.1-)
 DE (Hepaticin) (NS3p) (p70); Nonstructural protein 4A (NS4A) (p8);
 DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
 DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
 OS Hepatitis C virus (isolate Taiwan) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 ON NCBI_TaxID=31645;
 RN NUCLEOTIDE SEQUENCE [MENA]
 RP MEDLINE=92230206; PubMed=1314449;
 RA Chen P.-J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;
 RT "The Taiwanese hepatitis C virus genome: sequence determination and
 RL mapping the 5' termini of viral genomic and antigenomic RNA.";
 RL Virology 188:102-113(1992).
 [2]
 RP NUCLEOTIDE SEQUENCE OF 1460-1643.
 RC STRAIN=Isolate T3;
 RX PubMed=1648541;
 RA Chen P.-J., Lin M.H., Tu S.J., Chen D.S.;
 RT "Isolation of a complementary DNA fragment of hepatitis C virus in
 RL Taiwan revealed significant sequence variations compared with other
 RT isolates.";
 RL Hepatology 14:73-78(1991).
 [3]
 RP HOMOMULTIMERIZATION.
 RX MEDLINE=96207393; PubMed=8615040; DOI=10.1006/viro.1996.0164;
 RA Matsumoto M., Hsieh T.-Y., Jeng K.-S., Zhu N., Lai M.M.C.;
 RT "Homotypic interaction and multimerization of hepatitis C virus core
 RL protein.";
 RL Virology 218:43-51(1996).
 [4]
 RP INTERACTION OF CORE PROTEIN WITH HUMAN LTBR.
 RX PubMed=8995654;
 RA Matsumoto M., Hsieh T.-Y., Zhu N., VanArsdale T., Hwang S.B.,
 RA Jeng K.-S., Gorbaleya A.E., Lo S.-Y., Ou J.-H., Ware C.F.,
 RA Lai M.M.C.;
 RT "Hepatitis C virus core protein interacts with the cytoplasmic tail of
 RL lymphotxin-beta receptor.";
 RL J. Virol. 71:1301-1309(1997).
 [5]
 RP INTERACTION OF CORE PROTEIN WITH HUMAN TNFRSF1A.
 RX PubMed=9557650;
 RA Zhu N., Khoshnash A., Schneider R., Matsumoto M., Dennert G.,
 RA Ware C.F., Lai M.M.C.;
 RT "Hepatitis C virus core protein binds to the cytoplasmic domain of
 RT tumor necrosis factor (TNF) receptor 1 and enhances TNF-induced
 RT apoptosis.";
 RL J. Virol. 72:3691-3697(1998).
 [6]
 RP INTERACTION OF CORE PROTEIN WITH HUMAN HNRPK.
 RX MEDLINE=98316333; PubMed=9651361; DOI=10.1074/jbc.273.28.17651;
 RA Hsieh T.-Y., Matsumoto M., Chou H.-C., Schneider R., Hwang S.B.,
 RA Lee A.S., Lai M.M.C.;
 RT "Hepatitis C virus core protein interacts with heterogeneous nuclear
 RL ribonucleoprotein K.";
 RL J. Biol. Chem. 273:17651-17659(1998).
 [7]
 RP INTERACTION OF CORE PROTEIN WITH HUMAN APOA2.
 RX PubMed=10498661; DOI=10.1002/hep.510300429;
 RA Sabile A., Perlemuter G., Bono F., Kohara K., Demaugre F., Kohara M.,
 RA Matsuura Y., Miyamura T., Brechot C., Barba G.;
 RT "Hepatitis C virus core protein binds to apolipoprotein AII and its
 RT secretion is modulated by fibrates.";
 RL Hepatology 30:1064-1076(1999).
 [8]
 RP INTERACTION OF CORE PROTEIN WITH HUMAN P53.
 RX PubMed=10544138; DOI=10.1006/viro.1999.9979;
 RA Lu W., Lo S.-Y., Chen M., Wu K.-U., Fung Y.K.T., Ou J.-H.;
 RT "Activation of p53 tumor suppressor by hepatitis C virus core
 RT protein.";
 RL Virology 264:134-141(1999).

[9] INTERACTION OF CORE PROTEIN WITH HUMAN YWHAZ.
RA PubMed=16643444; DOI=10.1128/JVI.74.4.1736-1741.2000;
RA Aoki H., Hayashi J., Moriyama M., Arakawa Y., Hino O.;
RT "Hepatitis C virus core protein interacts with 14-3-3 protein and
RL activates the kinase Raf-1.";
RL J. Virol. 74:1736-1741(2000).
[10]
RN INTERACTION OF CORE PROTEIN WITH HUMAN CREB3.
RP PubMed=10675342; DOI=10.1093/emboj/19.4.729;
RA Jin D.-Y., Wang H.-L., Zhou Y., Chun A.C.S., Kibler K.V., Hou Y.-D.,
RA Kung H.-F., Jeang K.-T.;
RT "Hepatitis C virus core protein-induced loss of LZIP function
RL correlates with cellular transformation.";
RL EMBO J. 19:729-740(2000).
[11]
RN INTERACTION OF CORE PROTEIN WITH HUMAN STAT3.
RP MEDLINE=22198899; PubMed=12208879; DOI=10.1084/jem.20012127;
RA Yoshida T., Hanada T., Tokuhisa T., Kosei K.-I., Sata M., Kohara M.,
RA Yoshimura A.;
RT "Activation of STAT3 by the hepatitis C virus core protein leads to
RL cellular transformation.";
RL J. Exp. Med. 196:641-653(2002).
[12]
RN INTERACTION OF CORE PROTEIN WITH HUMAN CYTOKERATINS AND VIMENTIN.
RX PubMed=15846844; DOI=10.1002/jmcc.200401093;
RA Kang S.-W., Shin M.-J., Kim J.-H., Oh J.-W.;
RT "Proteomic profiling of cellular proteins interacting with the
RL hepatitis C virus core protein.";
RL Proteomics 5:2227-2237(2005).
[13]
RN FUNCTION OF CORE PROTEIN.
RX PubMed=15699147; DOI=10.1046/j.1365-2893.2000.00201.x;
RA Chou A.H., Tsai H.F., Wu Y.Y., Hu C.Y., Hwang L.H., Hsu P.I.,
RA Hsu P.N.;
RT "Hepatitis C virus core protein modulates TRAIL-mediated apoptosis by
RL enhancing Bid cleavage and activation of mitochondria apoptosis
RL signaling pathway.";
RL J. Immunol. 174:2160-2166(2005).
[14]
RN REVIEW.
RP PubMed=10718937; DOI=10.1046/j.1365-2893.2000.00201.x;
RA McLauchlan J.;
RT "Properties of the hepatitis C virus core protein: a structural
RL protein that modulates cellular processes.";
RL J. Viral Hepat. 7:2-14(2000).
[15]
RN REVIEW.
RP PubMed=14752815; DOI=10.1002/hep.20032;
RA Penin F., Dubuisson J., Rey F.A., Moradpour D., Pawlotsky J.-M.;
RT "Structural biology of hepatitis C virus.";
RL Hepatology 39:5-19(2004).
-1- FUNCTION: Core protein packages viral RNA to form a viral
CC nucleocapsid, and promotes virion budding. Modulates viral
CC translation initiation by interacting with HCV IRES and 40S
CC ribosomal subunit. Also regulates many host cellular functions
CC such as signaling pathways and apoptosis. Prevents the
CC establishment of cellular antiviral state by blocking the
CC interferon-alpha/beta (IFN-alpha/beta) and IFN-gamma signaling
CC pathways and by inducing human STAT1 degradation. Plays an
CC important role in virus-mediated cell transformation leading to
CC hepatocellular carcinomas. Interacts with, and activates SPAT3
CC leading to cellular transformation. May repress the promoter of
CC p53, and sequester CREB3 and Sp110 isoform 3/Sp110b in the
CC cytoplasm. Also represses cell cycle negative regulating factor
CC CDKN1A, thereby interrupting an important check point of normal
CC cell cycle regulation. Targets transcription factors involved in
CC the regulation of inflammatory responses and in the immune
CC response; suppresses NK-kappaB activation, and activates AP-1.
CC Mediates apoptotic pathways through association with TNF-type
CC receptors TNFRSF1A and LTRB, although its effect on death
CC receptors-induced apoptosis remains controversial. Enhances TRAIL
CC mediated apoptosis, suggesting that it might play a role in

CC immune-mediated liver cell injury. Secreted core protein is able
CC to bind ClqR1 at the T-cell surface, resulting in down-regulation
CC of T-lymphocytes proliferation. May transactivate human MYC, Rous
CC sarcoma virus LTR, and SV40 promoters. May suppress the human FOS
CC and HIV-1 LTR activity. May alter lipid metabolism by interacting
CC with hepatocellular proteins involved in lipid accumulation and
CC storage (By similarity).
CC -1- FUNCTION: Envelope glycoproteins E1 and E2 are involved in virus
CC attachment to the host cell as well as in virus endocytosis and
CC fusion with host membrane. E2 inhibits human PRKR activation,
CC preventing the establishment of an antiviral state (By
CC similarity).
CC -1- FUNCTION: p7 seems to be a hexameric ion channel protein
CC (viroporin) and is inhibited by the antiviral drug amantadine.
CC Also inhibited by long-alkyl-chain iminosugar derivatives.
CC Essential for infectivity (By similarity).
CC -1- FUNCTION: Protease NS2-3, which is a putative cysteine protease,
CC is responsible for the autocatalytic cleavage of NS2-NS3. Seems to
CC undergo self-inactivation following maturation (By similarity).
CC -1- FUNCTION: NS3 displays three distinct enzymatic activities: serine
CC protease, NTPase and RNA helicase. NS3 serine protease, in
CC association with NS4A, NS4B-NS5A and NS5A-NS5B. NS3/NS4A complex also
CC prevents phosphorylation of human Irf3, thus preventing the
CC establishment of dsRNA induced antiviral state. NS3 RNA helicase
CC binds to RNA and unwinds dsRNA in the 3' to 5' direction (By
CC similarity).
CC -1- FUNCTION: NS4B may induce a specific membrane alteration that
CC serves as a scaffold for the virus replication complex. This
CC membrane alteration gives rise to the so called ER-derived
CC membranous web that contains the replication complex (By
CC similarity).
CC -1- FUNCTION: NS5A is a component of the replication complex.
CC Downregulates viral IRES translation initiation. Mediates
CC interferon resistance, presumably by interacting with and
CC inhibiting human PRKR/PKR. The hyperphosphorylated form of NS5A is
CC an inhibitor of viral replication (By similarity).
CC -1- FUNCTION: NS5B is a RNA-dependent RNA polymerase that plays an
CC essential role in the virus replication (By similarity).
CC -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
CC precursor polyprotein, commonly with Asp or Glu in the P6
CC position, Cys or Thr in P1 and Ser or Ala in P1'.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -1- COFACTOR: Binds 1 zinc ion per NS3 protease domain. Binds 1 zinc
CC ion per NS5A N-terminal domain (By similarity).
CC -1- ENZYME REGULATION: Activity of protease NS2-3 is dependent on zinc

Query Match 10.28; Score 12; DB 1; Length 3009;
Best Local Similarity 100.0%; Pred. NO. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||
Db 1663 GGVLAALAAAYCL 1674

RESULT 156

Q59IP0_9HEPC
ID Q59IP0_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q59IP0;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DB Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=0;
RA Ikeda M., Abe K., Danesako H., Nakamura T., Naka K., Kato N.;

RT "Efficient replication of a full-length hepatitis C virus genome,
RT strain O, in cell culture, and development of a luciferase reporter
system.";
RL Biochem. Biophys. Res. Commun. 329:1350-1359 (2005).
DR EMBL; AB191333; BAD91386.1; -; Genomic_RNA.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326795 MW; A2529E4B86253DD6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAALAAAYCL 1675

RESULT 157
Q68285_9HEPC
ID Q68285_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q68285;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cho J., Park Y., Lee Y., Yang J., Ryu W.;
RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U16362; AAA52748.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; IXP.
DR SMR; Q68285; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.

DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
KW SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326915 MW; 29B306FC5B8EBC9E CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAALAAAYCL 1675

RESULT 158
Q68533_9HEPC
ID Q68533_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q68533;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Mueller H.M., Pfaff E., Heller A.E., Goesser T., Theilmann L.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U45476; AAA86907.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR HSP; Q81751; ICWX.
DR SMR; Q68533; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR011492; Flavi DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF07652; Flavi DEAD; 1.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.

DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327117 MW; 1D0B5A97C1466ED6 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||
DB 1664 GGVLAALAAAYCL 1675

RESULT 159

Q6GYR9_9HEPC
ID Q6GYR9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q6GYR9; 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=N589;
RX PubMed=15218169; DOI=10.1099/vir.0.79984-0;
RA Kalinina O., Norder H., Magnus L.O.;
RT "Full-length open reading frame of a recombinant hepatitis C virus
strain from St Petersburg: proposed mechanism for its formation.";
RL J. Gen. Virol. 85:1853-1857 (2004).
DR EMBL; AY587844; AAT00643.1; -; Genomic_RNA.
DR SMR; Q6GYR9; 1029-1657, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: translation; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR002868; HCV NS4b.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF02907; HCV NS2; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.

DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326545 MW; D05AB2897F142A58 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||
DB 1664 GGVLAALAAAYCL 1675

RESULT 160

Q807P3_9HEPC
ID Q807P3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q807P3; 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22047193; PubMed=12051758; DOI=10.1016/S0006-291X(02)00342-X;
RA Kishine H., Sugiyama K., Hijikata M., Kato N., Takahashi H., Noshi T.,
Nio Y., Hosaka M., Miyazaki Y., Shimotohno K.;
RT "Subgenomic replicon derived from a cell line infected with the
hepatitis C virus";
RL Biochem. Biophys. Res. Commun. 293:993-999 (2002).
DR EMBL; AB080299; BAC54896.1; -; Genomic_RNA.
DR SMR; Q807P3; 1029-1657, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: translation; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV RdRP.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF02907; HCV NS2; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF02907; HCV NS3; 1.


```
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; RDRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327097 MW; EE6418C7A723E686 CRC64;

Query Match          10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
   |||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 161
Q81825_9HEPC
ID Q81825_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q81825;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-NOV-2004 (TRENBLrel. 26, Last annotation update)
DE MRNA, complete cds.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Cho J.-M.;
RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; M96362; AA45721.1; -; mRNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; 1DXP.
DR SMR; Q81825; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR000745; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
```

```
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
SQ SEQUENCE 3010 AA; 326925 MW; FE997D54EE05142B CRC64;

Query Match          10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
   |||||
Db 1664 GGVLAALAAAYCL 1675

RESULT 162
Q8QRL8_9HEPC
ID Q8QRL8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q8QRL8;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yildiz E., Oztan A., Akkiz H., Ozturk M.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF483269; AAL91977.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q8QYSL; 1CWK.
DR SMR; Q8QRL8; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR000745; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
```

```
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptide.
SQ SEQUENCE 3010 AA; 326932 MW; 3B8CA9C861814CB CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 163
Q913V3_9HEPC
AC Q913V3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
NCBI_TaxID=11103;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14747471; DOI=10.1074/jbc.M312822200;
RA Teukiyama-Kohara K., Tone S., Maruyama I., Inoue K., Kataume A.,
RA Nuriya H., Ohmori H., Ohkawa J., Taira K., Hoshikawa Y., Shibasaki F.,
RA Reth M., Minatogawa Y., Kohara M.,
RT "Activation of the CKI-CDK-RB-E2F pathway in full genome hepatitis C
RT virus-expressing cells.";
RL J. Biol. Chem. 279:14531-14541(2004).
EMBL: AY045702; AAK97744.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q913V3; 1029-1657, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptide.
SQ SEQUENCE 3010 AA; 326939 MW; DS28A329A11F0D92 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 164
Q9DTE6_9HEPC
AC Q9DTE6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
NCBI_TaxID=11103;
RX [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
EMBL: AB049091; BAB18804.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; 1DXP.
DR SMR; Q9DTE6; 1029-1657, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
```

```
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR Polyprotein.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327043 MW; 3807DC6879684C95 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 165
Q9J3G8_9HBP
ID Q9J3G8_9HBP PRELIMINARY; PRT; 3010 AA.
AC Q9J3G8;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD24;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF207765; AAP65955.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWX.
DR SMR; Q9J3G8; 1029-1657; 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
```

DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327119 MW; E48938CA801B97AC CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

RESULT 167

Q9J3H3 9HEPC
ID Q9J3H3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MDL9;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.,
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RE EMBL; AF207760; AAF65950.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9J3H3; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F:serine-type peptidase activity; IEA.
DR GO; GO:0005138; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_P8.
DR InterPro; IPR007094; RNA_pol_P5vir.

DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327235 MW; 44C34677649CB8DD CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

RESULT 168

Q9QIX3 9HEPC
ID Q9QIX3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIX3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD9-2;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.,
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity";
RL Virology 263:244-253(1999).
DR EMBL; AF165062; AAD56197.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9QIX3; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.

```

DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_P8.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327254 MW; 9F1B0B3P536774FA CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 169
P88803_9HEPC
ID P88803_9HEPC PRELIMINARY; PRT; 3010 AA.
AC P88803;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-1b;
RX MEDLINE=95340824; PubMed=7542279;
RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
RA Yamamoto C., Izumi N., Marumo F., Sato C.;
RT "Comparison of full-length sequences of interferon-sensitive and
RT resistant hepatitis C virus 1b.";
RL J. Clin. Invest. 96:224-230(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-1b;
RA Enomoto N.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; D50484; BA009075.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR HSSP; P26663; 1NS3.
DR SMR; P88803; 1029-1657, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transmembrane; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.

```

```

DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_P8.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Envelope protein; Polyprotein.
FT CHAIN 1 191 core protein.
FT CHAIN 192 383 E1.
FT CHAIN 384 809 E2.
FT CHAIN 810 1026 NS2.
FT CHAIN 1027 1657 NS3.
FT CHAIN 1658 1711 NS4A.
FT CHAIN 1712 1972 NS4B.
FT CHAIN 1973 2419 NS5A.
FT CHAIN 2420 3010 NS5B.
SQ SEQUENCE 3010 AA; 327333 MW; 5F81505783F8F8B8 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 170
Q4PKP9_9HEPC
ID Q4PKP9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q4PKP9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yeh C.-T.;
RT "Emergence of mutation clusters in the HCV genome during sequential
RT viral passages in Sip-L expressing cells.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ071885; AY81920.1; -; mRNA.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327152 MW; 39DCAD60317840EC CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

```

Db 1664 GGVLAALAAAYCL 1675

RESULT 171

ID O09796 9HEPC PRELIMINARY; PRT; 3010 AA.
AC O09796;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Genomic RNA, complete cds.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=genotype II/1b;
RX MEDLINE=97170750; PubMed=9018054;
RA Sugiyama K., Kato N., Mizutani T., Ikeda M., Tanaka T., Shimotohno K.;
RT "Genetic analysis of the hepatitis C virus (HCV) genome from HCV-
RT infected human T cells.";
RL J. Gen. Virol. 78:329-336(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=9385694;
RA Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; D85516; BAA19625.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; O8JYSL; ICWX.
DR SMR; O09796; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_P8.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01538; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01536; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.

DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
SQ SEQUENCE 3010 AA; 327269 MW; 4324643E6A1D5CE4 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
|||||||
Db 1664 GGVLAALAAAYCL 1675
RESULT 172
O92969 9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92969;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St. Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL J. Virology 244:161-172(1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF054247; AAC15722.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; O02828; INB4.
DR SMR; O92969; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_P8.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01538; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01536; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.

```

DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326781 MW; 9B3FD910CF00E2C5 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCL 29
D 1664 GGVLAALAAYCL 1675

RESULT 173
O92970_9HEPC
ID O92970_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92970_
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR EMBL; AF054248; AAC15723.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q02828; I1N4.
DR SMR; Q92970; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.

```

```

DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RDRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326845 MW; 45PB399AD6141A88 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAYCL 29
D 1664 GGVLAALAAYCL 1675

RESULT 174
O92971_9HEPC
ID O92971_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92971_
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=HC-J4.
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St Clare M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence

```

```
RT J. Gen. Virol. 74:661-668 (1993).
RL BML; AF054249; AAC15724.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PQ0329; PQ0329.
DR HSP; Q02828; INB4.
DR SMR; Q92971; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326838 MW; 8FEB0CFB99CBF82 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 175
O92972_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O92972_9HEPC
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=HC-J4;
RC
```

```
RX MEDLINE=98240944; PubMed=9581788; DOI=10.1006/viro.1998.9092;
RA Yanagi M., St. Claire M., Shapiro M., Emerson S.U., Purcell R.H.,
RA Bukh J.;
RT "Transcripts of a chimeric clone of hepatitis C virus genotype 1b are
RT infectious in vivo.";
RL Virology 244:161-172 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR BML; AF054450; AAC15725.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0329; PQ0329.
DR HSP; Q02828; INB4.
DR SMR; Q92972; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326757 MW; A14CA7F4BB4197B5 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 176
O93016_9HEPC PRELIMINARY; PRT; 3010 AA.
ID O93016_9HEPC
AC O93016_9HEPC
```


DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=92044457; PubMed=1658209;
 RX Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 J. Gen. Virol. 72:2805-2809(1991).
 DR EMBL: AJ000009; CAA03854.1; -; Genomic_RNA.
 DR PIR: A61196; A61196.
 DR PIR: PQ0246; PQ0246.
 DR PIR: PQ0255; PQ0255.
 DR PIR: PS0329; PS0329.
 DR HSSP: Q8JYS1; 1CWK.
 DR SMR: Q93016; 1029-1657, 2008-2170, 2420-2949.
 DR GO: GO:0019028; C:viral capsid; IEA.
 DR GO: GO:0019031; C:viral envelope; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR GO: GO:0019079; P:viral genome replication; IEA.
 DR GO: GO:0019087; P:viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV core.
 DR InterPro: IPR002519; HCV env.
 DR InterPro: IPR002531; HCV NS1.
 DR InterPro: IPR000745; HCV NS4a.
 DR InterPro: IPR001490; HCV NS4b.
 DR InterPro: IPR002868; HCV NS5a.
 DR InterPro: IPR002166; HCV RdRp.
 DR InterPro: IPR001650; Helicase C.
 DR InterPro: IPR002518; Pept_039 HCV NS2.
 DR InterPro: IPR004109; Peptidase S2S.
 DR InterPro: IPR007095; RNA pol DS PS.
 DR InterPro: IPR007094; RNA pol_psvir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV core; 1.
 DR Pfam: PF01539; HCV env; 1.
 DR Pfam: PF01560; HCV NS1; 1.
 DR Pfam: PF01538; HCV NS2; 1.
 DR Pfam: PF02907; HCV NS3; 1.
 DR Pfam: PF01006; HCV NS4a; 1.
 DR Pfam: PF01001; HCV NS4b; 1.
 DR Pfam: PF01506; HCV NS5a; 1.
 DR Pfam: PF00271; Helicase C; 1.
 DR Pfam: PF00998; RGRP 3; 1.
 DR SMART: SM00487; DEXDC; 1.
 KW Polyprotein. 1 191 core protein.
 FT CHAIN 192 383 E1 protein.
 FT CHAIN 384 746 E2 protein.

FT CHAIN 747 809 p7 protein.
 FT CHAIN 810 1026 NS2 protein.
 FT CHAIN 1027 1657 NS3 protein.
 FT CHAIN 1658 1711 NS4a protein.
 FT CHAIN 1712 1972 NS4b protein.
 FT CHAIN 1973 2419 NS5a protein.
 FT CHAIN 2420 3010 NS5b protein.
 SQ SEQUENCE 3010 AA; 327124 MW; 0EE02EDAS4A8B61D CRC64;

 Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred.No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 1664 GGVLAALAAAYCL 1675

 RESULT 177
 O93077_9HEPC
 ID O93077_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC O93077;
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98122498; PubMed=9462666; DOI=10.1002/hep.510270242;
 RA Aizaki H., Aoki Y., Harada T., Ishii K., Suzuki T., Nagamori S.,
 RA Toda G., Matsuura Y., Miyamura T.;
 RT "Full-length complementary DNA of hepatitis C virus genome from an
 RT infectious blood sample.";
 J. Hepatology 27:621-627(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 J. Gen. Virol. 72:2805-2809(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93224886; PubMed=8385694;
 RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
 RA Dusheiko G., Saeed A.A., Holmes E.C.;
 RT "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity.";
 J. Gen. Virol. 74:661-668(1993).
 RL EMBL: D89815; BAA25076.1; -; Genomic_RNA.
 DR PIR: A61196; A61196.
 DR PIR: PQ0246; PQ0246.
 DR PIR: PQ0804; PQ0804.
 DR PIR: PS0329; PS0329.
 DR HSSP: Q8JYS1; 1CWK.
 DR SMR: Q93077; 1029-1657, 2008-2170, 2420-2949.
 DR GO: GO:0019028; C:viral capsid; IEA.
 DR GO: GO:0019031; C:viral envelope; IEA.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO: GO:0003723; F:RNA binding; IEA.
 DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: GO:0006350; P:transcription; IEA.
 DR GO: GO:0019079; P:viral genome replication; IEA.

DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327355 MW; 889BBA102A733390 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||
DB 1664 GGVLAALAAAYCL 1675

RESULT 178
P90191_9HEPC
ID P90191_9HEPC PRELIMINARY; PRT; 3010 AA.
AC P90191;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-1b;
RX MEDLINE=95340824; PubMed=7542279;
RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
RA Yamamoto C., Izumi N., Marumo F., Sato C.;
RT "Comparison of full-length sequences of interferon-sensitive and
RT resistant hepatitis C virus 1b.";
RL J. Clin. Invest. 96:224-230(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-1b;
RA Enomoto N.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagaaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";

J. Gen. Virol. 72:2805-2809(1991).
[4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; D50482; BAA09073.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR SMR; P90191; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Envelope protein; Polyprotein.
RN CHAIN 1 191 core protein.
FT CHAIN 192 383 E1.
FT CHAIN 384 809 E2.
FT CHAIN 810 1026 NS2.
FT CHAIN 1027 1657 NS3.
FT CHAIN 1658 1711 NS4A.
FT CHAIN 1712 1972 NS4B.
FT CHAIN 1973 2419 NS5A.
FT CHAIN 2420 3010 NS5B.
SQ SEQUENCE 3010 AA; 327442 MW; 5F15AC675A0C8268 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||
DB 1664 GGVLAALAAAYCL 1675

Db 1664 GGVLAALAAAYCL 1675

RESULT 179

P90193_9HEPC PRELIMINARY; PRT; 3010 AA.

AC P90193; 1

DT 01-MAY-1997 (TRENBLrel. 03, Created)

DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)

DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)

DE Polyprotein.

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=HCV-1b;

RX MEDLINE=95340824; PubMed=7542279;

RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,

RA Yamamoto C., Izumi N., Marumo F., Sato C.;

RT "Comparison of full-length sequences of interferon-sensitive and

RT resistant hepatitis C virus 1b.;"

RL J. Clin. Invest. 96:224-230(1995).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=HCV-1b;

RA Enomoto N.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=92044457; PubMed=1658209;

RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,

RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;

RT "cDNA clones of Japanese hepatitis C virus genomes derived from a

RT single patient show sequence heterogeneity.;"

RL J. Gen. Virol. 72:2805-2809(1991).

DR EMBL; D50485; BAA09076.1; -; Genomic_RNA.

DR PIR; A61196; A61196.

DR PIR; PQ0246; PQ0246.

DR PIR; PS0329; PS0329.

DR HSSP; Q8JYS1; 1CWK.

DR SMR; P90194; 1018-1657, 2008-2170, 2420-2949.

DR GO; GO:0019028; C:viral capsid; IEA.

DR PIR; P901543; HCV_capsid; 1.

DR PIR; P901542; HCV_core; 1.

DR PIR; P901539; HCV_env; 1.

DR Pfam; P901560; HCV_NS1; 1.

DR Pfam; P901538; HCV_NS2; 1.

DR Pfam; P902907; HCV_NS3; 1.

DR Pfam; P901006; HCV_NS4a; 1.

DR Pfam; P901001; HCV_NS4b; 1.

DR Pfam; P901506; HCV_NS5a; 1.

DR Pfam; P900271; Helicase_C; 1.

DR Pfam; P900998; RGRP_3; 1.

DR SMART; SM00487; DEXDc; 1.

KW Envelope protein; Polyprotein.

FT CHAIN 1 191 core protein.

FT CHAIN 192 383 E1.

FT CHAIN 384 809 E2.

FT CHAIN 810 1026 NS2.

FT CHAIN 1027 1657 NS3.

FT CHAIN 1658 1711 NS4A.

FT CHAIN 1712 1972 NS4B.

FT CHAIN 1973 2419 NS5A.

FT CHAIN 2420 3010 NS5B.

SQ SEQUENCE 3010 AA; 326889 MW; 21CD35B3DAC02B84 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29

|||||||

Db 1664 GGVLAALAAAYCL 1675

RESULT 180

P90194_9HEPC PRELIMINARY; PRT; 3010 AA.

AC P90194; 1

DT 01-MAY-1997 (TRENBLrel. 03, Created)

DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)

DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)

DE Polyprotein.

OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=11103;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=HCV-1b;

RX MEDLINE=95340824; PubMed=7542279;

RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,

RA Yamamoto C., Izumi N., Marumo F., Sato C.;

RT "Comparison of full-length sequences of interferon-sensitive and

RT resistant hepatitis C virus 1b.;"

RL J. Clin. Invest. 96:224-230(1995).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=HCV-1b;

RA Enomoto N.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=92044457; PubMed=1658209;

RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,

RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;

RT "cDNA clones of Japanese hepatitis C virus genomes derived from a

RT single patient show sequence heterogeneity.;"

RL J. Gen. Virol. 72:2805-2809(1991).

DR EMBL; D50481; BAA09072.1; -; Genomic_RNA.

DR PIR; A61196; A61196.

DR PIR; PQ0246; PQ0246.

DR PIR; PQ0254; PQ0254.

DR PIR; PS0329; PS0329.

DR HSSP; Q8JYS1; 1CWK.

DR SMR; P90194; 1018-1657, 2008-2170, 2420-2949.

DR GO; GO:0019028; C:viral capsid; IEA.

DR GO; GO:0019031; C:viral envelope; IEA.

DR GO: 0005524; F:ATP binding; IEA.
 DR GO: 0008026; F:ATP-dependent helicase activity; IEA.
 DR GO: 0003723; F:RNA binding; IEA.
 DR GO: 0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO: 0008236; F:serine-type peptidase activity; IEA.
 DR GO: 0005198; F:structural molecule activity; IEA.
 DR GO: 0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: 0006350; P:transcription; IEA.
 DR GO: 0019079; P:viral genome replication; IEA.
 DR GO: 0019087; P:viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR000745; HCV_NS4b.
 DR InterPro: IPR001490; HCV_NS4a.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRP.
 DR InterPro: IPR001650; Helicase C.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; RdRP 3; 1.
 DR Pfam: PF00271; Helicase C; 1.
 DR SMART: SM00487; DEXDc; 1.
 KW Envelope protein; Polyprotein.
 FT CHAIN 1 191 core protein.
 FT CHAIN 192 383 E1.
 FT CHAIN 384 809 E2.
 FT CHAIN 810 1026 NS2.
 FT CHAIN 1027 1657 NS3.
 FT CHAIN 1658 1711 NS4a.
 FT CHAIN 1712 1972 NS4b.
 FT CHAIN 1973 2419 NS5a.
 FT CHAIN 2420 3010 NS5b.
 SQ SEQUENCE 3010 AA; 326819 MW; 98D5C2A2D47FD011 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 1664 GGVLAALAAAYCL 1675

RESULT 181
 Q02828_9HEPC
 ID Q02828_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC Q02828;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DB Polypeptide.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_Taxid=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HC-J4;

XX MEDLINE=92391112; PubMed=1325713;
 RA Okamoto H., Kojima M., Okada S., Yoshizawa H., Iizuka H., Tanaka T.,
 RA Muchmore E.E., Peterson D.A., Ito Y., Mishihiro S.;
 RT "Genetic drift of hepatitis C virus during an 8.2-year infection in a
 RT chimpanzee: variability and stability.";
 RL Virology 190:894-899(1992).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=9204457; PubMed=1658209;
 RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 DR EMBL: D10750; BAA01583.1; -; Genomic_RNA.
 DR PIR: A61196; A61196.
 DR PIR: PQ0246; PQ0246.
 DR PIR: PS0329; PS0329.
 DR PDB: 1NB4; X-ray; A/B=2420-2989.
 DR PDB: 1NB6; X-ray; A/B=2420-2989.
 DR PDB: 1NB7; X-ray; A/B=2420-2989.
 DR SMR: Q02828; 1029-1657, 2008-2170.
 DR GO: 0019028; C:viral capsid; IEA.
 DR GO: 0019031; C:viral envelope; IEA.
 DR GO: 0005524; F:ATP binding; IEA.
 DR GO: 0008026; F:ATP-dependent helicase activity; IEA.
 DR GO: 0003723; F:RNA binding; IEA.
 DR GO: 0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO: 0008236; F:serine-type peptidase activity; IEA.
 DR GO: 0005198; F:structural molecule activity; IEA.
 DR GO: 0006508; P:proteolysis and peptidolysis; IEA.
 DR GO: 0006350; P:transcription; IEA.
 DR GO: 0019079; P:viral genome replication; IEA.
 DR GO: 0019087; P:viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH N.
 DR InterPro: IPR002522; HCV capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR000745; HCV_NS4b.
 DR InterPro: IPR001490; HCV_NS4a.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RdRP.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV capsid; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; RdRP 3; 1.
 DR SMART: SM00487; DEXDc; 1.
 KW Polyprotein.
 FT CHAIN 1 191 C protein.
 FT CHAIN 192 383 E protein.
 FT CHAIN 384 729 NS1/E2 protein.
 FT CHAIN 730 1006 NS2 protein.
 FT CHAIN 1007 1615 NS3 protein.
 FT CHAIN 1616 2013 NS4 protein.
 FT CHAIN 2014 3010 NS5 protein.
 SQ SEQUENCE 3010 AA; 326605 MW; F44C7FE7D288C242 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 182
Q02829_9HEPC Q02829_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q02829;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J4;
RA MEDLINE=92391112; PubMed=1325713;
RX Okamoto H., Kojima M., Okada S., Yoshizawa H., Iizuka H., Tanaka T.,
Rt Muchmore E.E., Peterson D.A., Ito Y., Mishiro S.;
RT "Genetic drift of hepatitis C virus during an 8.2-year infection in a
RT chimpanzee: variability and stability.";
RV Virology 190:894-899(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshina M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
Ra Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity";
RV J. Gen. Virol. 72:2805-2809(1991).
EN EMBL; D13558; BAA02756.1; -; Genomic_RNA.
DR FIR; A61196; A61196.
DR FIR; PQ0246; PQ0246.
DR FIR; PS0329; PS0329.
DR HSSP; QBYSI; LCWX.
SNR; Q02829; 1029-1657, 2008-2170, 2420-2949.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
GO; GO:0003723; F:RNA binding; IEA.
GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F:serine-type peptidase activity; IEA.
GO; GO:0006508; F:structural molecule activity; IEA.
GO; GO:0006350; P:transcription; IEA.
GO; GO:0019079; P:vital genome replication; IEA.
GO; GO:0019087; P:vital transformation; IEA.
InterPro; IPRO01410; DEAD.
InterPro; IPRO11545; DEAD/DEAH_N.
InterPro; IPRO02522; HCV capsid.
InterPro; IPRO02521; HCV core.
InterPro; IPRO02519; HCV env.
InterPro; IPRO02531; HCV NS1.
InterPro; IPRO00745; HCV NS4a.
InterPro; IPRO01490; HCV NS4b.
InterPro; IPRO02868; HCV NS5a.
InterPro; IPRO02166; HCV RdRp.
InterPro; IPRO02518; Pept U39 HCV NS2.
InterPro; IPRO04109; Peptidease S29.
InterPro; IPRO07095; RNA_pol_DS_PS.
InterPro; IPRO07094; RNA_pol_PSVir.
Pfam; PF01543; HCV capsid; 1.
Pfam; PF01542; HCV core; 1.
Pfam; PF01539; HCV env; 1.
Pfam; PF01560; HCV NS1; 1.
Pfam; PF01538; HCV NS2; 1.
Pfam; PF02907; HCV NS3; 1.
Pfam; PF01006; HCV NS4a; 1.

DR	Pfam; PF01001; HCV_NS4b; 1.	
DR	Pfam; PF01506; HCV_NS5a; 1.	
DR	Pfam; PF00998; RDRP_3; 1.	
DR	SMART; SM00487; DEXDc; 1.	
DR	Polyprotein.	
FT	CHAIN 1 191	C protein.
FT	CHAIN 192 383	E protein.
FT	CHAIN 384 729	NS1/E2 protein.
FT	CHAIN 730 1006	NS2 protein.
FT	CHAIN 1007 1615	NS3 protein.
FT	CHAIN 1616 2013	NS4 protein.
FT	CHAIN 2014 3010	NS5 protein.
SQ	SEQUENCE 3010 AA; 325962 MW; 93D465526F3EADP8 CRC64;	

Query Match	10.2%;	Score 12;	DB 2;	Length 3010;
Best Local Similarity	100.0%;	Pred. No. 0.032;		
Matches 12;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	18	GGVLAALAAAYCL 29	
Db	1664	GGVLAALAAAYCL 1675	

RESULT 183		
QSR2A8_9HEPC		
ID	QSR2A8_9HEPC PRELIMINARY;	PRT; 3010 AA.
AC	QSR2A8;	
DT	01-FEB-2005 (trEMBLrel. 29, Created)	
DT	01-FEB-2005 (trEMBLrel. 29, Last sequence update)	
DT	01-FEB-2005 (trEMBLrel. 29, Last annotation update)	
DS	Polyprotein (Fragment).	
OS	Hepatitis C virus type 1b.	
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	
OC	Hepacivirus.	
OX	NCBI_TaxID=31647;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RC	STRAIN=No. 31;	
RA	Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,	
RA	Kenny-Walsh E., Shanahan F., Watanabe M.;	
RT	"HCV genotype 1b full-length sequence 1-31.";	
RL	Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AB154206; BAD73999.1; -; Genomic RNA.	
DR	SNR; QSR2A8; 1029-1657, 2008-2170, 2420-2949.	
DR	CG; GO:0016021; C:integral to membrane; IEA.	
DR	CG; GO:0019028; C:viral capsid; IEA.	
DR	CG; GO:0019031; C:viral envelope; IEA.	
DR	CG; GO:0005524; F:ATP binding; IEA.	
DR	CG; GO:0008026; F:ATP-dependent helicase activity; IEA.	
DR	CG; GO:0016787; F:hydrolase activity; IEA.	
DR	CG; GO:0003723; F:RNA binding; IEA.	
DR	CG; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.	
DR	CG; GO:0008236; F:serine-type peptidase activity; IEA.	
DR	CG; GO:0005198; F:structural molecule activity; IEA.	
DR	CG; GO:0016740; F:proteinase activity; IEA.	
DR	CG; GO:0006508; P:proteolysis and peptidolysis; IEA.	
DR	CG; GO:0006350; P:transcription; IEA.	
DR	CG; GO:0019079; P:viral genome replication; IEA.	
DR	CG; GO:0019087; P:viral transformation; IEA.	
DR	InterPro; IPR001410; DEAD.	
DR	InterPro; IPR011545; DEAD/DEAH N.	
DR	InterPro; IPR002522; HCV capsid.	
DR	InterPro; IPR002521; HCV core.	
DR	InterPro; IPR002519; HCV env.	
DR	InterPro; IPR002531; HCV NS1.	
DR	InterPro; IPR000745; HCV NS4a.	
DR	InterPro; IPR001490; HCV NS4b.	
DR	InterPro; IPR002868; HCV NS5a.	
DR	InterPro; IPR002166; HCV RdRp.	
DR	InterPro; IPR001650; Helicase_C.	
DR	InterPro; IPR002518; Pept U39 HCV NS2.	
DR	InterPro; IPR004109; Peptidase S29.	
DR	InterPro; IPR007095; RNA pol DS PS.	

```
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF01001; HCV_NS4a; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 327438 MW; 5B2063DD2C51B249 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 184
QSR2A9_9HEPC
ID QSR2A9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2A9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 30;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154205; BAD73998.1; -; Genomic RNA.
DR SMR; QSR2A9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003236; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformatation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
```

```
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PSVir.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 327411 MW; 0CDB61CE741A89D9 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 185
QSR2B0_9HEPC
ID QSR2B0_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 29;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154204; BAD73997.1; -; Genomic RNA.
DR SMR; QSR2B0; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003236; F:RNA binding; IEA.
DR GO; GO:0003236; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019087; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR001490; HCV_NS4a.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR000745; HCV_NS4b.
```

```

DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW NON_TER
FT NON_TER
SQ SEQUENCE 3010 3010
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 186
QSR2B1_9HEPC
ID QSR2B1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=31647;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 28;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DSC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154203; BAD73996.1; -; Genomic RNA.
DR SMR; QSR2B1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0008026; P: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; P: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 187
QSR2B2_9HEPC
ID QSR2B2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=31647;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 27;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154202; BAD73995.1; -; Genomic RNA.
DR SMR; QSR2B2; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0008026; P: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; P: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH_N.

```

```

DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
DR Polyprotein.
KW NON_TER
FT
SQ SEQUENCE 3010 AA; 327202 MW; 62291219B217AA12 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 188
QSR2B3_9HEPC
ID QSR2B3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 26;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154201; BAD73994.1; -; Genomic RNA.
DR SMR; QSR2B3; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005224; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.

```

```

DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
DR Polyprotein.
KW NON_TER
FT
SQ SEQUENCE 3010 AA; 327324 MW; 78C019D87E5AFF76 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 189
QSR2B4_9HEPC
ID QSR2B4_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 25;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154200; BAD73993.1; -; Genomic RNA.
DR SMR; QSR2B4; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005224; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.

```


DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006740; F:transferase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV core.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR002518; Pept U39_HCV NS2.
 DR InterPro; IPR004109; Peptidase_S23.
 DR InterPro; IPR007094; RNA pol_PSVir.
 DR Pfam; PF01543; HCV capsid; 1.
 DR Pfam; PF01542; HCV core; 1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV NS1; 1.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RdRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICC; 1.
 KW Polyprotein.
 FT NON TER 3010 3010
 SQ SEQUENCE 3010 AA; 327440 MW; ED490B979C93C6F7 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred.No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 1664 GGVLAALAAAYCL 1675

RESULT 190
 QSR2B5_9HEPC
 AC QSR2B5;
 DT 01-FEB-2005 (TRENBLrel. 29, Created)
 DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis C virus type 1b.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=31647;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=No. 24;
 RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
 RA Kenny-Walsh E., Shanahan F., Watanabe M.;
 RT "HCV genotype 1b full-length sequence 1-31.";
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB154199; BAD73992.1; -; Genomic RNA.
 DR SMR; QSR2B5; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006740; F:transferase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH_N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV core.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR002518; Pept U39_HCV NS2.
 DR InterPro; IPR004109; Peptidase_S23.
 DR InterPro; IPR007094; RNA pol_PSVir.
 DR Pfam; PF01543; HCV capsid; 1.
 DR Pfam; PF01542; HCV core; 1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV NS1; 1.
 DR Pfam; PF01538; HCV NS2; 1.
 DR Pfam; PF02907; HCV NS3; 1.
 DR Pfam; PF01006; HCV NS4a; 1.
 DR Pfam; PF01001; HCV NS4b; 1.
 DR Pfam; PF01506; HCV NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RdRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICC; 1.
 KW Polyprotein.
 FT NON TER 3010 3010
 SQ SEQUENCE 3010 AA; 327102 MW; 796CE22CCA467849 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred.No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
 |||||
 Db 1664 GGVLAALAAAYCL 1675

RESULT 191
 QSR2B6_9HEPC
 ID QSR2B6_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC QSR2B6;
 DT 01-FEB-2005 (TRENBLrel. 29, Created)
 DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis C virus type 1b.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=31647;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=No. 23;
 RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
 RA Kenny-Walsh E., Shanahan F., Watanabe M.;
 RT "HCV genotype 1b full-length sequence 1-31.";
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB154198; BAD73991.1; -; Genomic RNA.
 DR SMR; QSR2B6; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005524; F: ATP binding; IEA.
DR GO: 0008026; F: ATP-dependent helicase activity; IEA.
DR GO: 0016787; F: hydrolase activity; IEA.
DR GO: 0003723; F: RNA binding; IEA.
DR GO: 0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO: 0008236; F: serine-type peptidase activity; IEA.
DR GO: 0016740; F: structural molecule activity; IEA.
DR GO: 0003723; F: RNA binding; IEA.
DR GO: 0006508; F: proteolysis and peptidolysis; IEA.
DR GO: 0006350; P: transcription; IEA.
DR GO: 0019079; P: viral genome replication; IEA.
DR GO: 0019087; P: viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR02519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR002868; HCV NS4b.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept_U39_HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRP 3; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 346893 MW; FBCF578A68B25957 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

RESULT 192
QSR2B7_9HEPC
ID QSR2B7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 22;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.,
RT "HCV genotype 1b full-length sequence 1-31.";

RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB154197; BAD73990.1; -: Genomic RNA.
DR SMR: QSR2B7.1029-1657, 2008-2170, 2420-2949.
DR GO: 0016021; C: integral to membrane; IEA.
DR GO: 0019028; C: viral capsid; IEA.
DR GO: 0019031; C: viral envelope; IEA.
DR GO: 0005524; F: ATP binding; IEA.
DR GO: 0008026; F: ATP-dependent helicase activity; IEA.
DR GO: 0016787; F: hydrolase activity; IEA.
DR GO: 0003723; F: RNA binding; IEA.
DR GO: 0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO: 0008236; F: serine-type peptidase activity; IEA.
DR GO: 0005198; F: structural molecule activity; IEA.
DR GO: 0016740; F: transferase activity; IEA.
DR GO: 0006508; P: proteolysis and peptidolysis; IEA.
DR GO: 0006350; P: transcription; IEA.
DR GO: 0019079; P: viral genome replication; IEA.
DR GO: 0019087; P: viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH_N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR02519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept_U39_HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRP 3; 1.
DR SMART: SM00487; DEXDC; 1.
DR SMART: SM00490; HELIC; 1.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327237 MW; 9097312599348D9A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

RESULT 193
QSR2B8_9HEPC
ID QSR2B8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2B8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
[1]
RP NUCLEOTIDE SEQUENCE.

```
RC STRAIN-No. 21;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154196; BAD73989.1; -; Genomic RNA.
DR SMR; Q5R2B8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV env.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR007095; RNA pol_D5_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR SMART; SM00998; RdRp_3; 1.
DR Polyprotein.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 065BBF77DEBFD9C2 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||
Db 1664 GGVLAALAAAYCL 1675

RESULT 194
Q5R2B9_9HEPC
ID Q5R2B9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q5R2B9;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=31647;
```

```
RN STRAIN-No. 20;
RC Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154195; BAD73988.1; -; Genomic RNA.
DR SMR; Q5R2B9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019031; C:viral capsid; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV env.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR SMART; SM00998; RdRp_3; 1.
DR Polyprotein.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 1505FF2812180BF5B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||
Db 1664 GGVLAALAAAYCL 1675

RESULT 195
Q5R2C0_9HEPC
ID Q5R2C0_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q5R2C0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
```

OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-NO. 19;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154194; BAD73987.1; -; Genomic RNA.
DR SMR; QSR2C0; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005224; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0008236; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; F: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR001666; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Peptidase_S29.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007094; RNA_pol_DS_PS.
DR InterPro; IPR007095; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 373737 MW; 1B56341E22FF756 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAAALAAAYCL 1675

RESULT 196
QSR2C1_9HEPC
ID QSR2C1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-NO. 18;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154193; BAD73986.1; -; Genomic RNA.
DR SMR; QSR2C1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005224; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0008236; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR001666; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Peptidase_S29.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007094; RNA_pol_DS_PS.
DR InterPro; IPR007095; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 327064 MW; 3A46E79935416A02 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAAALAAAYCL 1675

RESULT 197
QSR2C2_9HEPC
ID QSR2C2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C2;

01-FEB-2005 (TrEMBLrel. 29, Created)
01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
Polyprotein (Fragment)
Hepatitis C virus type 1b.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepadnavirus.
NCBI_TaxID=31647;
[1]
NUCLEOTIDE SEQUENCE.
STRAIN=No. 17;
Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
Kenny-Walsh E., Shanahan F., Watanabe M.;
"HCV genotype 1b full-length sequence 1-31.";
Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AB154192; BAD73985.1; -; Genomic RNA.
SMR; 05R2C3; 1029-1657, 2008-2170, 2420-2949.
GO; GO:0016021; C: integral to membrane; IEA.
GO; GO:0005524; F: ATP binding; IEA.
GO; GO:0016787; F: hydrolase activity; IEA.
GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
GO; GO:0019031; C: viral envelope; IEA.
GO; GO:0003723; F: RNA binding; IEA.
GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F: serine-type peptidase activity; IEA.
GO; GO:0005198; F: structural molecule activity; IEA.
GO; GO:0016740; P: protease activity; IEA.
GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
GO; GO:0006350; P: transcription; IEA.
GO; GO:0019079; P: viral genome replication; IEA.
GO; GO:0019087; P: viral transformation; IEA.
InterPro; IPR001410; DEAD.
InterPro; IPR011545; DEAD/DEAH N.
InterPro; IPR002522; HCV capsid.
InterPro; IPR002531; HCV env.
InterPro; IPR002519; HCV core.
InterPro; IPR002531; HCV NS1.
InterPro; IPR000745; HCV NS4a.
InterPro; IPR001490; HCV NS4b.
InterPro; IPR002868; HCV NS5a.
InterPro; IPR002166; HCV RdRP.
InterPro; IPR001650; Helicase C.
InterPro; IPR002518; Pept U39 HCV NS2.
InterPro; IPR004109; Peptidase S29.
InterPro; IPR007095; RNA pol DS PS.
InterPro; IPR007094; RNA pol_PSVir.
Pfam; PF01543; HCV capsid; 1.
Pfam; PF01542; HCV core; 1.
Pfam; PF01539; HCV env; 1.
Pfam; PF01560; HCV NS1; 1.
Pfam; PF01538; HCV NS2; 1.
Pfam; PF02907; HCV NS3; 1.
Pfam; PF01006; HCV NS4a; 1.
Pfam; PF01001; HCV NS4b; 1.
Pfam; PF01506; HCV NS5a; 1.
Pfam; PF00271; Helicase C; 1.
Pfam; PF00998; RdRP 3; 1.
SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELICc; 1.
Polyprotein.
NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327770 MW; 087FDE5DC2114DA3 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29

|||||
Db 1664 GGVLAAALAAAYCL 1675

RESULT 198
05R2C3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC 05R2C3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
DE Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepadnavirus.
OC NCBI_TaxID=31647;
[1]
NUCLEOTIDE SEQUENCE.
STRAIN=No. 16;
Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
Kenny-Walsh E., Shanahan F., Watanabe M.;
"HCV genotype 1b full-length sequence 1-31.";
Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AB154191; BAD73984.1; -; Genomic RNA.
SMR; 05R2C3; 1029-1657, 2008-2170, 2420-2949.
GO; GO:0016021; C: integral to membrane; IEA.
GO; GO:0019028; C: viral capsid; IEA.
GO; GO:0019031; C: viral envelope; IEA.
GO; GO:0005524; F: ATP binding; IEA.
GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
GO; GO:0016787; F: hydrolase activity; IEA.
GO; GO:0003723; F: RNA binding; IEA.
GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F: serine-type peptidase activity; IEA.
GO; GO:0005198; F: structural molecule activity; IEA.
GO; GO:0016740; P: transferase activity; IEA.
GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
GO; GO:0006350; P: transcription; IEA.
GO; GO:0019079; P: viral genome replication; IEA.
GO; GO:0019087; P: viral transformation; IEA.
InterPro; IPR001410; DEAD.
InterPro; IPR011545; DEAD/DEAH N.
InterPro; IPR002522; HCV capsid.
InterPro; IPR002521; HCV core.
InterPro; IPR002519; HCV env.
InterPro; IPR002531; HCV NS1.
InterPro; IPR000745; HCV NS4a.
InterPro; IPR001490; HCV NS4b.
InterPro; IPR002868; HCV NS5a.
InterPro; IPR002166; HCV RdRP.
InterPro; IPR001650; Helicase C.
InterPro; IPR002518; Pept U39 HCV NS2.
InterPro; IPR004109; Peptidase S29.
InterPro; IPR007095; RNA pol DS PS.
InterPro; IPR007094; RNA pol_PSVir.
Pfam; PF01543; HCV capsid; 1.
Pfam; PF01542; HCV core; 1.
Pfam; PF01539; HCV env; 1.
Pfam; PF01560; HCV NS1; 1.
Pfam; PF01538; HCV NS2; 1.
Pfam; PF02907; HCV NS3; 1.
Pfam; PF01006; HCV NS4a; 1.
Pfam; PF01001; HCV NS4b; 1.
Pfam; PF01506; HCV NS5a; 1.
Pfam; PF00271; Helicase C; 1.
Pfam; PF00998; RdRP 3; 1.
SMART; SM00487; DEXDC; 1.
SMART; SM00490; HELICc; 1.
Polyprotein.
NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327350 MW; 17E4A7DDA2EDDC1B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29

```
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 199
QSR2C4_9HEPC
ID QSR2C4_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP STRAIN=No. 15;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AB154190; BAD73983.1; -; Genomic RNA.
SMR; QSR2C4; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR SMART; SM00487; RdRp 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR Polyprotein.
KW NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327248 MW; 49BA895C093E80ED CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
```

```
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327320 MW; 9D93F03A57ABCOFF CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAALAAAYCL 1675

RESULT 201
QSR2C6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C6;
DT 01-FEB-2005 (TRENBLrel. 29, Created)
DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 13;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154187; BAD73981.1; -; Genomic RNA.
DR SMR; QSR2C6; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR002518; Pept U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF02907; HCV NS2; 1.
DR Pfam; PF01538; HCV NS3; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.

RESULT 202
QSR2C7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C7;
DT 01-FEB-2005 (TRENBLrel. 29, Created)
DT 01-FEB-2005 (TRENBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 12;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154187; BAD73980.1; -; Genomic RNA.
DR SMR; QSR2C7; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR002518; Pept U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF02907; HCV NS2; 1.
DR Pfam; PF01538; HCV NS3; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
```

```
DR Pfam; PF00998; RDRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327559 MW; 3BAD51D5B9BA44F1 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||
Db 1664 GGVLAALAAAYCL 1675

RESULT 203
QSR2C8_9HEPC
ID QSR2C8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C8_9HEPC
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RC STRAIN-No. 11;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AB154186; BAD73979.1; -; Genomic RNA.
DR SMR; QSR2C8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0016787; F:ATP binding; IEA.
DR GO; GO:000524; F:RNA binding; IEA.
DR GO; GO:001787; F:hydrolase activity; IEA.
DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0016740; P:transferrase activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
```

```
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RDRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327195 MW; 7AF6FB8CA2EF997 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||
Db 1664 GGVLAALAAAYCL 1675

RESULT 204
QSR2C9_9HEPC
ID QSR2C9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2C9_9HEPC
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RC STRAIN-No. 10;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
EMBL; AB154185; BAD73978.1; -; Genomic RNA.
DR SMR; QSR2C9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:000524; F:ATP binding; IEA.
DR GO; GO:0016787; F:ATP-dependent helicase activity; IEA.
DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferrase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
```



```
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR SMART; SM00490; HELICc; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 327210 MW; A985B065015BD2D7 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 205
QSR2D0_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2D0;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RC STRAIN=No. 9;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154184; BAD73977.1; -; Genomic RNA.
DR SMR; QSR2D0; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR00745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RGRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 206
QSR2D1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2D1;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RC STRAIN=No. 8;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154183; BAD73976.1; -; Genomic RNA.
DR SMR; QSR2D1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR00745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RGRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
```

```
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 327268 MW; FCPAED39228B841A CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

RESULT 207
QSR2D2_9HEPC
ID QSR2D2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2D2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 7;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154182; BAD73975.1; -; Genomic RNA.
DR SMR; QSR2D2; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0008236; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
```

```
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
KW Polyprotein.
FT NON_TER 3010 3010
SQ SEQUENCE 3010 AA; 327397 MW; A545EA731FB5908C CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

RESULT 208
QSR2D3_9HEPC
ID QSR2D3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2D3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 6;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB154181; BAD73974.1; -; Genomic RNA.
DR SMR; QSR2D3; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0008236; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
```

DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept U39 HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA pol DS PS.
DR InterPro: IPR007094; RNA pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327170 MW; 52BCBB701AEFD333 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 1664 GGVLAALAAAYCL 1675

RESULT 209
QSR2D4_9HEPC
ID QSR2D4_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2D4;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 5;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB154180; BAD73973.1; -; Genomic RNA.
DR QSR2D4; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0016740; F:transfrase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 1664 GGVLAALAAAYCL 1675

RESULT 210
QSR2D5_9HEPC
ID QSR2D5_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSR2D5;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=No. 4;
RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,
RA Kenny-Walsh E., Shanahan F., Watanabe M.;
RT "HCV genotype 1b full-length sequence 1-31.";
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB154179; BAD73972.1; -; Genomic RNA.
DR QSR2D5; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0016740; F:transfrase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.

GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 GO; GO:0006350; P:transcription; IEA.
 GO; GO:0019079; P:viral genome replication; IEA.
 GO; GO:0019087; P:viral transformation; IEA.

InterPro; IPR001410; DEAD.
 InterPro; IPR011545; DEAD/DEAH N.
 InterPro; IPR002522; HCV capsid.
 InterPro; IPR002519; HCV env.
 InterPro; IPR002531; HCV NS1.
 InterPro; IPR000745; HCV NS4a.
 InterPro; IPR001490; HCV NS4b.
 InterPro; IPR002868; HCV NS5a.
 InterPro; IPR002166; HCV RdRp.
 InterPro; IPR001650; Helicase C.
 InterPro; IPR002518; Pept_U39_HCV NS2.
 InterPro; IPR004109; Peptidase S29.
 InterPro; IPR007095; RNA_pol_DS_PS.
 InterPro; IPR007094; RNA_pol_PSVir.

Pfam; PF01543; HCV capsid; 1.
 Pfam; PF01539; HCV core; 1.
 Pfam; PF01560; HCV NS1; 1.
 Pfam; PF01538; HCV NS2; 1.
 Pfam; PF02907; HCV NS3; 1.
 Pfam; PF01006; HCV NS4a; 1.
 Pfam; PF01001; HCV NS4b; 1.
 Pfam; PF01506; HCV NS5a; 1.
 Pfam; PF00271; Helicase C; 1.
 Pfam; PF00998; RdRp 3; 1.
 SMART; SM00487; DEXDc; 1.
 Polyprotein.

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 18 GGVLAAALAYCL 29

DB 1664 GGVLAAALAYCL 1675

RESULT 211

QSR2D6_9HEPC

ID QSR2D6_9HEPC PRELIMINARY; PRT; 3010 AA.

AC QSR2D6;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)

DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis C virus type 1b.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=31647;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=NO. 3;

RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,

RA Kenny-Walsh E., Shanahan F., Watanabe M.;

RT "HCV genotype 1b full-length sequence 1-31.";

RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.

EMBL; AB154178; BAD73971.1; -; Genomic RNA.

SMR; QSR2D6; 1029-1657, 2008-2170, 2420-2949.

GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0019028; C:viral capsid; IEA.

GO; GO:0019031; C:viral envelope; IEA.

GO; GO:0005244; P:ATP binding; IEA.

GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

GO; GO:0016787; F:hydrolase activity; IEA.

GO; GO:0003723; F:RNA binding; IEA.

GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.

GO; GO:0008236; P:serine-type peptidase activity; IEA.
 GO; GO:0005198; P:structural molecule activity; IEA.
 GO; GO:0016740; F:transferase activity; IEA.
 GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 GO; GO:0006350; P:transcription; IEA.

GO; GO:0019079; P:viral genome replication; IEA.
 GO; GO:0019087; P:viral transformation; IEA.

InterPro; IPR001410; DEAD.

InterPro; IPR011545; DEAD/DEAH N.

InterPro; IPR002522; HCV capsid.

InterPro; IPR002519; HCV core.

InterPro; IPR002531; HCV env.

InterPro; IPR002531; HCV NS1.

InterPro; IPR000745; HCV NS4a.

InterPro; IPR001490; HCV NS4b.

InterPro; IPR002868; HCV NS5a.

InterPro; IPR002166; HCV RdRp.

InterPro; IPR001650; Helicase C.

InterPro; IPR002518; Pept_U39_HCV NS2.

InterPro; IPR004109; Peptidase S29.

InterPro; IPR007095; RNA_pol_DS_PS.

InterPro; IPR007094; RNA_pol_PSVir.

Pfam; PF01543; HCV capsid; 1.

Pfam; PF01539; HCV core; 1.

Pfam; PF01560; HCV NS1; 1.

Pfam; PF01538; HCV NS2; 1.

Pfam; PF02907; HCV NS3; 1.

Pfam; PF01006; HCV NS4a; 1.

Pfam; PF01001; HCV NS4b; 1.

Pfam; PF01506; HCV NS5a; 1.

Pfam; PF00271; Helicase C; 1.

Pfam; PF00998; RdRp 3; 1.

SMART; SM00487; DEXDc; 1.

Polyprotein.

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query 18 GGVLAAALAYCL 29

DB 1664 GGVLAAALAYCL 1675

RESULT 212

QSR2D7_9HEPC

ID QSR2D7_9HEPC PRELIMINARY; PRT; 3010 AA.

AC QSR2D7;

DT 01-FEB-2005 (TrEMBLrel. 29, Created)

DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis C virus type 1b.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.

OX NCBI_TaxID=31647;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=NO. 2;

RA Itakura J., Nagayama K., Enomoto N., Hamano K., Fanning L.J.,

RA Kenny-Walsh E., Shanahan F., Watanabe M.;

RT "HCV genotype 1b full-length sequence 1-31.";

RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.

EMBL; AB154177; BAD73970.1; -; Genomic RNA.

SMR; QSR2D7; 1029-1657, 2008-2170, 2420-2949.

GO; GO:0016021; C:integral to membrane; IEA.

GO; GO:0019028; C:viral capsid; IEA.

GO; GO:0019031; C:viral envelope; IEA.

GO; GO:0005244; P:ATP binding; IEA.

DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0016787; F:hydrolyase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept U39 HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
FT NON TER 3010 3010
SQ SEQUENCE 3010 AA; 327346 MW; B06ED3E1205EE02 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAALAAAYCL 1675

RESULT 213
Q68788_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q68788;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE HCV polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96362158; PubMed=8720135;
RA Seki M., Honda Y.;
RT "Phosphorothioate antisense oligodeoxynucleotides capable of
RT inhibiting hepatitis C virus gene expression: in vitro translation
RT assay.";
RL J. Biochem. 118:1199-1204 (1995).
RS [2]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
ET "cDNA clones of Japanese hepatitis C virus genomes derived from a
ET single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR EMBL; D45172; BAA08120.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; IDEXP.
DR SMR; Q68788; 1029-1657, 2008-2170, 2420-2949.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept U39 HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326884 MW; EED840E6A050E766 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAALAAAYCL 1675

RESULT 214
Q68826_9HEPC PRELIMINARY; PRT; 3010 AA.
ID Q68826_9HEPC
AC Q68826;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Cho M.J.;
 RC STRAIN=J33;
 RL Submitted (SEP-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
 RA Dushenko G., Saeed A.A., Holmes E.C.;
 RT "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity.";
 RL J. Gen. Virol. 74:661-668(1993).
 DR EMBL; D14484; BAA03375.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PQ0804; PQ0804.
 DR PIR; PS0329; PS0329.
 DR HSP; O8JY81; 1CW.
 DR SMR; Q68826; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV core.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR001650; Helicase C.
 DR Polyprotein.
 SQ SEQUENCE 3010 AA; 327154 MW; 7270F47984554PAD CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred.No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCL 29
 |||||
 Db 1664 GGVLAALAAAYCL 1675
 RESULT 215
 Q68833_9HEPC
 ID Q68833_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC Q68833;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HCV-L2;
 RA Cho J.M., Park Y.-W., Lee Y.-B., Yang J.-Y., Kim C.-H., Choo S.-H.,
 RA Ryu W.-S.;
 RT "Molecular cloning of hepatitis C virus genome from a single Korean
 RT blood donor.";
 RL Mol. Cells 5:317-324(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HCV-L2;
 RA Cho J.;
 RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 DR EMBL; U01214; AAA75355.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PS0329; PS0329.
 DR HSP; O8JY81; 1CW.
 DR SMR; Q68833; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002521; HCV core.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR000745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR001650; Helicase C.
 DR Polyprotein.

DR	GO; GO:0019028;	C:viral capsid; IEA.
DR	GO; GO:0019031;	C:viral envelope; IEA.
DR	GO; GO:0005524;	F:ATP binding; IEA.
DR	GO; GO:0008026;	P:ATP-dependent helicase activity; IEA.
DR	GO; GO:0003723;	F:RNA binding; IEA.
DR	GO; GO:0003968;	F:RNA-directed RNA polymerase activity; IEA.
DR	GO; GO:0008236;	F:serine-type peptidase activity; IEA.
DR	GO; GO:0005198;	F:structural molecule activity; IEA.
DR	GO; GO:0006508;	P:proteolysis and peptidolysis; IEA.
DR	GO; GO:0006350;	P:transcription; IEA.
DR	GO; GO:0019079;	P:viral genome replication; IEA.
DR	GO; GO:0019087;	P:viral transformation; IEA.
DR	InterPro; IPR001410;	DEAD.
DR	InterPro; IPR011545;	DEAD/DEAH_N.
DR	InterPro; IPR002522;	HCV capsid.
DR	InterPro; IPR002521;	HCV_core.
DR	InterPro; IPR002519;	HCV env.
DR	InterPro; IPR002531;	HCV NS1.
DR	InterPro; IPR000745;	HCV_NS4a.
DR	InterPro; IPR001490;	HCV_NS4b.
DR	InterPro; IPR002868;	HCV_NS5a.
DR	InterPro; IPR002166;	HCV_RdRP.
DR	InterPro; IPR001650;	Helicase_C.
DR	InterPro; IPR002518;	Pept_U39 HCV NS2.
DR	InterPro; IPR004109;	Peptidase_S29.
DR	InterPro; IPR007095;	RNA_pol_DS_PS.
DR	InterPro; IPR007094;	RNA_pol_pstvir.
DR	Pfam; PF01543;	HCV_capsid; 1.
DR	Pfam; PF01542;	HCV_core; 1.
DR	Pfam; PF01539;	HCV_env; 1.
DR	Pfam; PF01560;	HCV_NS1; 1.
DR	Pfam; PF01538;	HCV_NS2; 1.
DR	Pfam; PF02907;	HCV_NS3; 1.
DR	Pfam; PF02907;	HCV_NS4a; 1.
DR	Pfam; PF01006;	HCV_NS4b; 1.
DR	Pfam; PF01001;	HCV_NS4d; 1.
DR	Pfam; PF01506;	HCV_NS5a; 1.
DR	Pfam; PF00271;	Helicase_C; 1.
DR	Pfam; PF00998;	RdRP_3; 1.
DR	SMART; SM00487;	DENDC; 1.
FT	CHAIN	1 191
FT	CHAIN	192 383
FT	CHAIN	384 733
FT	CHAIN	734 1006
FT	CHAIN	1007 1615
FT	CHAIN	1616 1862
FT	CHAIN	1863 2013
FT	CHAIN	2014 3010
SQ	SEQUENCE	3010 AA; 327182 MW; 451740D8FF6134DA CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;		
Best Local Similarity 100.0%; Pred.No. 0.032;		
Matches 12; Conservative 0; Mismatches 0; Indels		
Qy	18 GGVLAALAAAYCL 29	
Db	1664 GGVLAALAAAYCL 1675	
RESULT 217		
ID Q81541.9HEPC		
DB Q81541.9HEPC PRELIMINARY;		
AC Q81541;		
DT 01-NOV-1996 (TREMblrel. 01, Created)		
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)		
DE 01-MAR-2004 (TREMblrel. 26, Last annotation update)		
DE Polyprotein.		
OS Hepatitis C virus.		
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.		
OX NCBI_TaxID=11103;		
ON [1]		
RP NUCLEOTIDE SEQUENCE.		
RC STRAIN=JT;		

XX MEDLINE=92259714; PubMed=1318627; DOI=10.1016/0168-1702(92)90066-1;
RA Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J., Nakazawa T.,
RA Hijikata M., Ishimura Y., Shimotohno K.;
RT "Molecular cloning of hepatitis C virus genome from a single Japanese
carrier: sequence variation within the same individual and among
infected individuals.";
RL Virus Res. 23:39-53(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; D11355; BAA18894.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; P80329; P80329.
DR HSP; Q81755; IDXP.
DR SMR; Q81541; 1034-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Envelope protein; Polyprotein.
FT CHAIN 1 191 put. p22.
FT CHAIN 192 383 put. p23.
FT CHAIN 384 729 put. p24.
FT CHAIN 730 1006 put. NS2.
FT CHAIN 1007 1615 put. NS3.
FT CHAIN 1616 1862 put. NS4a.
FT CHAIN 1863 2013 put. NS4b.
FT CHAIN 2014 3010 put. NS5.
SQ SEQUENCE 3010 AA; 326572 MW; 05F0B2102CF9DD9D CRC64;

Query Match

10.2%; Score 12; DB 2; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 218
Q81757_9HEPC
ID Q81757_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q81757;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; D30613; BAA06303.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; P80329; P80329.
DR HSP; Q81755; IDXP.
DR SMR; Q81757; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Envelope protein; Polyprotein.
FT CHAIN 1 191 put. p22.
FT CHAIN 192 383 put. p23.
FT CHAIN 384 729 put. p24.
FT CHAIN 730 1006 put. NS2.
FT CHAIN 1007 1615 put. NS3.
FT CHAIN 1616 1862 put. NS4a.
FT CHAIN 1863 2013 put. NS4b.
FT CHAIN 2014 3010 put. NS5.
SQ SEQUENCE 3010 AA; 326572 MW; 05F0B2102CF9DD9D CRC64;


```

DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RDRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326823 MW; 49C7CAGE22468523 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 219
Q81760_9HEPC
ID Q81760_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q81760;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8394876;
RA Wang Y., Okamoto H., Tenda F., Nagayama R., Tao Q.M., Mishiro S.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX Simmonds P., McOmish F., Yap F.L., Chan S.-W.W., Lin C.K.,
RA Dubeiko G., Saeed A.A., Holmes E.C.;
RA "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; D10934; BAA01728.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q81760; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.

```

```

DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326860 MW; EAYD306A4BA2E224 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 220
Q81989_9HEPC
ID Q81989_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q81989;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE E1 and E2/NS1 envelope glycoprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-N;
RA Zheng W.Z.;
RT "Genotype identification of hepatitis C virus (HCV) isolated from a
RT single Japanese carrier in Nagasaki prefecture and genome analysis of
RT E1 and E2/NS1 envelope glycoprotein regions.";
RL Nippon Netta Igakkai Zasshi 22:169-177(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-N;
RA Zheng W.-Y.Z.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";

```

J. Gen. Virol. 72:2805-2809(1991).
EMBL; D63857; BAA09919.1; -; Genomic_RNA.
DR PIR; PQ0246; PQ0246.
DR HSP; Q81755; IDXP.
DR SNR; Q81989; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV NS1.
DR InterPro; IPR002531; HCV NS4a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV env; 1.
DR Pfam; PF01539; HCV core; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXdc; 1.
KW Envelope protein.
SQ SEQUENCE 3010 AA; 327507 MW; C7BDB38169D6E3CF CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 221
Q8V638_9HEPC
ID Q8V638_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q8V638;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN NUCLEOTIDE SEQUENCE.
RP Fanning L.J., Itakura J., Nagayama K., Enomoto N.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658203;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
[3]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=93224886; PubMed=8385694;
RX Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
virus: identification of a new virus type and restrictions on sequence
diversity.";
J. Gen. Virol. 74:661-668(1993).
RL EMBL; AF313916; AAL55821.1; -; mRNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; ICW.
DR SNR; Q8V638; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV env; 1.
DR Pfam; PF01539; HCV core; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXdc; 1.
KW Envelope protein.
SQ SEQUENCE 3010 AA; 327507 MW; C7BDB38169D6E3CF CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 221
Q8V638_9HEPC
ID Q8V638_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q8V638;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN NUCLEOTIDE SEQUENCE.
RP Fanning L.J., Itakura J., Nagayama K., Enomoto N.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
[2]

```
Db      1664 GGVLAALAAAYCL 1675
|||||
RESULT 222
Q99AU2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q99AU2_9HEPC PRELIMINARY; PRT; 3010 AA.
AT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus type 1b.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=31647;
RN [1]
RP STRAIN=chimera of HCV-BK;
RX MEDLINE=21534507; PubMed=11677216;
RA Thomson M., Nascimbeni M., Gonzales S., Murthy K.K., Rehmann B.,
RA Liang T.J.;
RT "Emergence of a distinct pattern of viral mutations in chimpanzees
RT infected with a homogeneous inoculum of hepatitis C virus.";
RL Gastroenterology 121:1226-1233(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=9385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF333324; AAK08509.1; -; mRNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR SMR; Q99AU2; 1029-1657, 2008-2170.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR002518; PpT U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_poi_DS_PS.
```

```
DR InterPro; IPR007094; RNA_poi_PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEAD; 1.
DR Polyprotein.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327011 MW; 053B9A653B0AB335 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
|||||

RESULT 223
Q9DTD6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTD6_9HEPC PRELIMINARY; PRT; 3010 AA.
AT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AB049101; BAB18814.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; O8JYSL; 1CW.
DR SMR; Q9DTD6; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
```

```

DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327112 MW; DE182D810EF78EE4 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 224
Q9D7D7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9D7D7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
[2]
RP NUCLEOTIDE SEQUENCE.
RC MSDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AB049100; BAB18813.1; -; Genomic_RNA.
DR FNR; A61196; A61196.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR H8SP; Q8JYS1; 1CWX.
DR SMR; Q9D7D7; 1029-1657, 2008-2170, 2420-2949.

```

```

DR MER09S; S29_002; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326913 MW; 5505C62EB0DA0519 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 225
Q9D7D9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9D7D9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).

```

```

RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AB049098; BAB18811.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0329; PQ0329.
DR HSP; P26684; IHE1.
DR SMR; Q9D7D9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR0011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.

SQ SEQUENCE 3010 AA; 326599 MW; 8EA6737401DECFE CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 226
Q9DTE1_9HEPC
ID Q9DTE1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTE1_
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

```

DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RdRP_3; 1.
 DR SMART; SM00487; DEXDc; 1.
 KW Polyprotein.
 SQ SEQUENCE 3010 AA; 326848 MW; 90457AC819A32150 CRC64;
 Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCL 29
 DB 1664 GGVLAALAAAYCL 1675
 RESULT 227
 Q9DTE2_9HEPC
 ID Q9DTE2_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC Q9DTE2;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Serum;
 RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
 RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
 RA Mishihiro S.;
 RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
 RT with hepatocellular carcinoma: the 'progression score' revisited.";
 RL Hepatol. Res. 20:161-171(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Serum;
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=93224886; PubMed=8385694;
 RA Simmonds P., McOmish F., Yap F.L., Chan S.-W.W., Lin C.K.,
 RA Dushenko G., Saeed A.A., Holmes E.C.;
 RT "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity.";
 RL J. Gen. Virol. 74:661-668(1993).
 DR EMBL; AB049095; BAB18808.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PQ0252; PQ0252.
 DR PIR; PQ0253; PQ0253.
 DR PIR; PQ0254; PQ0254.
 DR PIR; PQ0255; PQ0255.
 DR PIR; PQ0804; PQ0804.
 DR PIR; PS0329; PS0329.
 DR HSP; Q8JY51; 1CWV.
 DR SRS; Q9DTE2; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.

DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral genome transformation; IEA.
 DR InterPro; IPR001410; DEAD
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV_capsid.
 DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NS1.
 DR InterPro; IPR000745; HCV_NS4a.
 DR InterPro; IPR001490; HCV_NS4b.
 DR InterPro; IPR002868; HCV_NS5a.
 DR InterPro; IPR002166; HCV_RdRP.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RdRP_3; 1.
 DR SMART; SM00487; DEXDc; 1.
 KW Polyprotein.
 SQ SEQUENCE 3010 AA; 326871 MW; 25BES4B9D7EEAA15 CRC64;
 Query Match 10.2%; Score 12; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred. No. 0.032;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 18 GGVLAALAAAYCL 29
 DB 1664 GGVLAALAAAYCL 1675
 RESULT 228
 Q9DTE4_9HEPC
 ID Q9DTE4_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC Q9DTE4;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Serum;
 RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
 RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
 RA Mishihiro S.;
 RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
 RT with hepatocellular carcinoma: the 'progression score' revisited.";
 RL Hepatol. Res. 20:161-171(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=93224886; PubMed=8385694;
 RA Simmonds P., McOmish F., Yap F.L., Chan S.-W.W., Lin C.K.,
 RA Dushenko G., Saeed A.A., Holmes E.C.;
 RT "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity.";
 RL J. Gen. Virol. 74:661-668(1993).
 DR EMBL; AB049095; BAB18808.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PQ0252; PQ0252.
 DR PIR; PQ0253; PQ0253.
 DR PIR; PQ0254; PQ0254.
 DR PIR; PQ0255; PQ0255.
 DR PIR; PQ0804; PQ0804.
 DR PIR; PS0329; PS0329.
 DR HSP; Q8JY51; 1CWV.
 DR SRS; Q9DTE2; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.

```
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish P., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dumeheille G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AB049093; BAB18806.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9DT84; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 327328 MW; 3DE6CF249BD1151C CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 229
QSDT85 9HEPC
ID QSDT85_9HEPC PRELIMINARY; PRT; 3010 AA.
AC QSDT85;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
```

```
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 230
Q9DTE7_9HEPC
ID Q9DTE7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTE7;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
RN [2]
RP MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yegasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese Hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AB049090; BAB18803.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00252; P00252.
DR PIR; P00253; P00253.
DR PIR; P00254; P00254.
DR PIR; P00255; P00255.
DR PIR; P00329; P00329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9DTE7; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019087; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5b.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
```

```
InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326841 MW; 58EE3BD4140BB588 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 231
Q9DTE8_9HEPC
ID Q9DTE8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTE8;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
RN [2]
RP MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AB049089; BAB18802.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; P00804; P00804.
DR PIR; P00329; P00329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9DTE8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
```



```
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327071 MW; E26F4D669A836C80 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 232
Q9DTE9_9HEPC
ID Q9DTE9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTE9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=111103;
RN [1]_TaxID=111103;
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
```

```
RT InterPro; IPR001410; DEAD.
RT InterPro; IPR011545; DEAD/DEAH N.
RT InterPro; IPR002522; HCV capsid.
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AB049088; BAB18801.1; -, Genomic_RNA.
DR F01; A61196; A61196.
DR F01; PQ0246; PQ0246.
DR F01; PQ0804; PQ0804.
DR F01; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q9DTE9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR00410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326783 MW; 668CFEAF5FEC3658 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 233
Q9DTE9_9HEPC
ID Q9DTE9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9DTE9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
```

```
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishihiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma; the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tauchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AB049087; BAB18800.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0254; PQ0254.
DR PIR; PS0329; PS0329.
DR HSP; Q8JVS1; 1CWX.
DR SMR; Q8JVS1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F:RNA-binding; IEA.
DR GO; GO:0006508; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Peptidase_S25.
DR InterPro; IPR004109; Peptidase_S25.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327156 MW; 7A9C7B1273266FF3 CRC64;

Query March 10.2k; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0k; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RGP3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327128 MW; 25324D4A7AAAD83B CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps
QY 18 GGVLAALAAAYCL 29
|||
DB 1664 GGVLAALAAAYCL 1675
RESULT 236
Q9J3G1_9HEPC
ID Q9J3G1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC ASJ3G1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD31;
RA Nagayama K., Kurotaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshina M., Tsuchiya M., Yagasaki M., Orita T., Haegawa M.,
RA Tomonoh K., Koijima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity";
RL J. Gen. Virol. 74:661-668(1993).
RL EMBL; AF207772; AAF65962.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR HSP; Q87YS1; ICWX.
DR SMR; Q9J3G1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; P:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

```

DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 327134 MW; FE2292587F421874 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 237
Q9J3G2_9HEPC
ID Q9J3G2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3G2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=MD30;
RC Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese Hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR EMBL; AF207771; AAF65961.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; P02446; P02446.
DR PIR; P02552; P02552.
DR PIR; P02553; P02553.
DR PIR; P02554; P02554.

```

```

DR PIR; P02555; P02555.
DR PIR; P0329; P0329.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q9J3G2; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326434 MW; C91E2BFAAA07D3E0 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 238
Q9J3G3_9HEPC
ID Q9J3G3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3G3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=MD29;
RC Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]

```

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
R Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
ET single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207770; AAF65960.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JY51; 1CW.
DR SMR; QJ3G3; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327233 MW; 19783535650CACE3 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 239
QJ3G4_9HEPC PRELIMINARY; PRT; 3010 AA.
ID QJ3G4;
AC QJ3G4;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA STRAIN=MD28;
RC Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
R Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
ET single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207769; AAF65959.1; -; Genomic_RNA.
DR PIR; PQ0246; PQ0246.
DR HSSP; Q8JY51; 1CW.
DR SMR; QJ3G4; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327322 MW; 0595728AC62464F3 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 240
QJ3G5_9HEPC PRELIMINARY; PRT; 3010 AA.
ID QJ3G5;
AC QJ3G5;


```

DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327169 MW; 74FAB6B80F24837B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 242
O9J3G7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O9J3G7_
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxid=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD25;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
[3]
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9224886; PubMed=8385594;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dubeiko G., Saeed A.A., Holmes E.C.;
RA "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF207766; AAF65956.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0255; PQ0255.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JTS1; 1CWX.
DR SMR; Q9J3G7; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.

```

```

DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326511 MW; 9EAC8ADA4B441DF CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 243
O9J3G9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O9J3G9_
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxid=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD23;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207764; AAF65954.1; -; Genomic_RNA.

```

```

DR PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00246; P00246.
DR HSP; Q8JY51; 1CW.
DR SMR; Q9J3H0; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0008026; F: RNA binding; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0003350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002521; HCV capsid.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
DR Polyprotein.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327486 MW; E87E3B0C5092F5A2 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 244
Q9J3H0_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RX [1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN-MD22;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.

```

```

RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RT Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207763; AAF65953.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; P00246; P00246.
DR PIR; P00254; P00254.
DR PIR; P00329; P00329.
DR HSP; Q8JY51; 1CW.
DR SMR; Q9J3H0; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C: viral capsid; IEA.
DR GO; GO:0019031; C: viral envelope; IEA.
DR GO; GO:0005524; P: ATP binding; IEA.
DR GO; GO:0008026; F: RNA binding; IEA.
DR GO; GO:0003723; F: RNA binding; IEA.
DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F: serine-type peptidase activity; IEA.
DR GO; GO:0005198; F: structural molecule activity; IEA.
DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P: transcription; IEA.
DR GO; GO:0019079; P: viral genome replication; IEA.
DR GO; GO:0019087; P: viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR SMART; SM00487; DEXDC; 1.
DR Polyprotein.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327101 MW; 737EEF31E3C2B28D CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675
RESULT 245
Q9J3H2_9HEPC PRELIMINARY; PRT; 3010 AA.
ID Q9J3H2_9HEPC PRELIMINARY;
AC Q9J3H2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```


DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD20;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207751; AAF65951.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q8JVS1; 1CW.
DR SMR; Q8J3H2; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005234; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326767 MW; 1A48EE4BE51440D0 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCL 29
| | | | | | | | | |
Db 1664 GGVLAAALAAAYCL 1675

RESULT 246
Q9J3H5_9HEPC PRELIMINARY; PRT; 3010 AA.
ID Q9J3H5;
AC Q9J3H5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD17;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207758; AAF65948.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PS0329; PS0329.
DR HSP; Q8JVS1; 1CW.
DR SMR; Q8J3H5; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005234; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.

KW Polyprotein.
SQ SEQUENCE 3010 AA; 326806 MW; 9FEE3D1B93B7AA4B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 1664 GGVLAALAAAYCL 1675

RESULT 247
Q9J3H6_9HEPC
ID Q9J3H6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD16;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tauchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF207757; AAF65947.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0255; PQ0255.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q9J3H6; 1029-1657, 2008-2170, 2420-2949.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
GO; GO:0003723; F:RNA binding; IEA.
GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F:serine-type peptidase activity; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
GO; GO:0006350; P:transcription; IEA.
GO; GO:0019079; P:viral genome replication; IEA.
GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.

DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S25.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327115 MW; DDF4162DA6969512 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 1664 GGVLAALAAAYCL 1675

RESULT 248
Q9J3H7_9HEPC
ID Q9J3H7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OC Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD15;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tauchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF207756; AAF65946.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q9J3H7; 1029-1657, 2008-2170, 2420-2949.

DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR000745; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S23.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327370 MW; D8653F7317FFA106 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAALAAAYCL 29
|||||
DB 1664 GGVLAAALAAAYCL 1675
RESULT 249
ID Q9J3H8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H8_9HEPC PRELIMINARY; PRT; 3010 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
single patient show sequence heterogeneity.";
single patient show sequence heterogeneity.";

J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207755; AAF65945.1; -; Genomic_RNA.
DR PIR; P0246; A61196.
DR PIR; P0246; A61196.
DR PIR; P0255; P0255.
DR PIR; P0255; P0255.
DR PIR; P0329; P0329.
DR PIR; P0329; P0329.
DR SMSP; Q81755; LDXP.
DR HSR; Q9J3H8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR000745; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S23.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326425 MW; B109B6487CD206E8 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAAALAAAYCL 29
|||||
DB 1664 GGVLAAALAAAYCL 1675
RESULT 250
ID Q9J3H9_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J3H9_9HEPC PRELIMINARY; PRT; 3010 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```
RC STRAIN=MD13;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9204457; PubMed=1658209;
RA Oshima M., Tauchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207754; AAF65944.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q9J3H9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:000826; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR00745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326989 MW; AF12CC00E0A8B078 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 251
Q9J310_9HEPC
ID Q9J310_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J310;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OC Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD12;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9204457; PubMed=1658209;
RA Oshima M., Tauchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207753; AAF65943.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q9J310; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:000826; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326696 MW; 074098DB305AF1A9 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
```

Db 1664 GGVLAALAAAYCL 1675
|||||
RESULT 252
Q9J311_9HEPC
ID Q9J311_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9J311;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=MD11;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF207752; AAF65942.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9J311; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent RNA polymerase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS5a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Rept_u39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.

DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327197 MW; F88BA81174E19444 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675
RESULT 253
Q9QIX1_9HEPC
ID Q9QIX1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIX1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=MD10-2;
RX NUCLEOTIDE SEQUENCE.
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
J. Virol. 72:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dubeiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF165064; AAD56199.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9QIX1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent RNA polymerase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS5a.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Rept_u39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.

```
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF02907; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326906 MW; 04233B5981E71EDD CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 254
O9QIX2_9HEPC
ID O9QIX2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O9QIX2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD10-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Haegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
```

```
J. Gen. Virol. 74:661-668(1993).
[4]
NUCLEOTIDE SEQUENCE.
RX PubMed=1314471;
RA Kato N., Ootsuyama Y., Tanaka T., Nakagawa M., Nakazawa T.,
RA Muraiso K., Ohkoshi S., Hijikata M., Shimotohno K.;
RT "Marked sequence diversity in the putative envelope proteins of
RT hepatitis C viruses.";
RL Virus Res. 22:107-123(1992).
DR EMBL; AF165063; AAD56198.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; P80329; P80329.
DR PIR; S24086; S24086.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q9QIX2; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326860 MW; 3806A4AF819ED552 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 255
O9QIX4_9HEPC
ID O9QIX4_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O9QIX4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
```

DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD9-1;
RA MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR EMBL; AF165061; AAD56196.1; -; Genomic_RNA.
DR FIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q8JYS1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Peptidase S29.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol Pevir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327391 MW; 94C94662C4A47695 CRC64;

Query Match

10.2%; Score 12; DB 2; Length 3010;

Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675
RESULT 256
SQ Q8JYS1 9HEPC
ID Q8JYS1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q8JYS1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD8-2;
RA MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity";
RL Virology 263:244-253 (1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD8-2;
RA MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity";
RL J. Gen. Virol. 72:2805-2809 (1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=9324886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity";
RL J. Gen. Virol. 74:661-668 (1993).
DR EMBL; AF165060; AAD56195.1; -; Genomic_RNA.
DR FIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q8JYS1; 1029-1657, 2008-2170, 2420-2949.
DR MEROPS; S29.002; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Peptidase S29.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol Pevir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327391 MW; 94C94662C4A47695 CRC64;

```
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept U39 HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF01538; HCV NS3; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRp 3; 1.
DR SMART: SM00487; DEXDC; 1.
DR Polyprotein.
KW POLYPEPTIDE.
SQ SEQUENCE 3010 AA; 327298 MW; 8B99F1EBA6A50F56 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 257
Q9QIX6_9HEPC
ID Q9QIX6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIX6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD8-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tarawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Koichiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusehko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
RL EMBL; AF165059; AAD56194.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
```

```
DR PIR: PQ0246; PQ0246.
DR PIR: PQ0254; PQ0254.
DR PIR: PQ0804; PQ0804.
DR PIR: PS0329; PS0329.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q9JX6; 1029-1657, 2008-2170, 2420-2949.
DR MEROPS; S29_002; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS1.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR001650; Helicase C.
DR InterPro: IPR002518; Pept U39 HCV NS2.
DR InterPro: IPR004109; Peptidase S29.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase C; 1.
DR Pfam: PF00998; RdRp 3; 1.
DR SMART: SM00487; DEXDC; 1.
DR Polyprotein.
KW POLYPEPTIDE.
SQ SEQUENCE 3010 AA; 327471 MW; 4613744EC4D4A013 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 258
Q9QIX7_9HEPC
ID Q9QIX7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIX7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_Taxid=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD7-2;
```


EX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF165058; AAD56193.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; 1CWX.
DR SMR; Q9QIX7; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 326978 MW; A3556D74F0C3AD2B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 18 GGVLAALAAAYCL 29
|||||
Db 1664 GGVLAALAAAYCL 1675
RESULT 259
Q9QIX8_9HEPC
ID Q9QIX8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIX8;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD7-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF165057; AAD56192.1; -, Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; 1CWX.
DR SMR; Q9QIX8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR Polyprotein.

```

DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326990 MW; 55F505A208C6B5CD CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 260
Q9QIY1_9HEPC
ID Q9QIY1_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=MD5-2;
RX MEDLINE=9204457; PubMed=1658209;
RA Oshima M., Tauchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9224885; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF165054; AAS56189.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HBBP; Q8JVS1; 1CWX.

```

```

DR SMR; Q9QIY1; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008366; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326724 MW; DE61BDF27B7AF3E0 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 261
Q9QIY2_9HEPC
ID Q9QIY2_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RC STRAIN=MD5-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).

```

RA NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX EMBL; AF165053; AAD56188.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q9QIY2; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR007095; RNA pol PS p8.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01003; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptide.
SQ SEQUENCE 3010 AA; 326599 MW; D186BA7E92F0B5E8 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
18 GGVLAALAAAYCL 29
|||||||

Db 1664 GGVLAALAAAYCL 1675
RESULT 262
Q9QIY3_9HEPC
ID Q9QIY3_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD4-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX EMBL; AF165052; AAD56187.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q9QIY3; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR007095; RNA pol PS p8.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01003; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptide.
SQ SEQUENCE 3010 AA; 326599 MW; D186BA7E92F0B5E8 CRC64;

```

DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_core; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 327092 MW; 2PF78321696D4002 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 263
Q9QIY4_9HEPC
ID Q9QIY4_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9204457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AF165051; AAD56186.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWV.
DR SMR; Q9QIY4; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.

```

```

DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
KW POLYPEPTIDE.
SQ SEQUENCE 3010 AA; 326834 MW; F6559AB2CFC3CB2 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 264
Q9QIY5_9HEPC
ID Q9QIY5_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
DE Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD3-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9204457; PubMed=1658209;

```

RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF165050; AAD56185.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q9QIY5; 2-112, 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 327434 MW; 15190E3463DE8C35 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675
RESULT 265
Q9QIY6_9HEPC
ID Q9QIY6_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

OK NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD3-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Negayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF165049; AAD56184.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q9QIY6; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_Ps.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3010 AA; 327372 MW; 998C7F293EABEC8D CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAAALAAAYCL 29
Db 1664 GGVLAAALAAAYCL 1675

```
RESULT 266
O9QIY7_9HEPC
ID O9QIY7_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O9QIY7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD2-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RL EMBL; AF165048; AAD56193.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0254; PQ0254.
DR PIR; P80329; P80329.
DR HSP; Q8JY81; ICWX.
DR SMR; Q9QIY7; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol ds ps.
DR InterPro; IPR007094; RNA pol BSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.

DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326422 MW; 060FA943F3FCAA8C CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 267
O9QIY8_9HEPC
ID O9QIY8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O9QIY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD2-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RL EMBL; AF165047; AAD56182.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0254; PQ0254.
DR PIR; P80329; P80329.
DR HSP; Q8JY81; ICWX.
DR SMR; Q9QIY8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol ds ps.
DR InterPro; IPR007094; RNA pol BSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
```

```
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326627 MW; 9952890F46183217 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 268
Q9QIY9 SHEPC
ID Q9QIY9_SHEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QIY9;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD1-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
DR EMBL; AF165046; AAD56181.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0253; PQ0253.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0255; PQ0255.
DR PIR; PS0329; PS0329.
DR HSSP; Q81755; 1DXP.

SMR; Q9QIY9; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326370 MW; D816D3BBBF14EE46 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
DB 1664 GGVLAALAAAYCL 1675

RESULT 269
Q9QI20 SHEPC
ID Q9QI20_SHEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QI20;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD1-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
RA Tazawa J.I., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
```



```
FT CHAIN 384 746 glycoprotein E2.
FT CHAIN 747 809 p7 peptide.
FT CHAIN 810 1026 NS2 proteinase.
FT CHAIN 1027 1657 NS3 proteinase/helicase.
FT CHAIN 1658 1712 NS3/4A proteinase cofactor.
FT CHAIN 1712 1972 NS4B protein.
FT CHAIN 1973 2419 NS5A phosphoprotein.
FT CHAIN 2419 3010 NS5B RNA dependant RNA polymerase.
SQ SEQUENCE 3010 AA; 327003 MW; A570BB980DD64634 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCL 29
Db 1664 GGVLAALAAYCL 1675

RESULT 271
Q9QP61_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9QP61;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=274933RU;
RA Mokhonov V.V., Samokhvalov E.I., Novikov D.V., Shatalov A.G.,
RA Prilipov A.G.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR PIR: A61196; A61196.
DR PIR: PQ0246; PQ0246.
DR PIR: FS0329; FS0329.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q9QP61; 1029-1657, 2008-2170, 2420-2949.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; P:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; P:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; F:viral genome replication; IEA.
DR GO: GO:0019087; P:viral genome replication; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
```

```
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S23.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRp 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 327072 MW; 9105F69483DD5BBA CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAYCL 29
Db 1664 GGVLAALAAYCL 1675

RESULT 272
Q9WIK8_9HEPC PRELIMINARY; PRT; 3010 AA.
AC Q9WIK8;
DT 01-NOV-1999 (TREMELrel. 12, Created)
DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98033183; PubMed=9367360;
RA Rispeter K., Lu M., Lechner S., Zibert A.;
RT "Cloning and characterization of a complete open reading frame of the
RT hepatitis C virus genome in only two cDNA fragments.";
RL J. Gen. Virol. 78:2751-2759 (1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Rispeter K.;
RL Thesis (1998), Universitaetsklinikum Essen, Institut fuer Virologie.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668 (1993).
```

```
DR EMBL; AJ132997; CAB41951.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CWX.
DR SMR; Q9WIK8; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 356877 MW; 0EC653BA29F8BAFE CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 273
Q9D7D8_9HEPC
ID Q9D7D8_9HEPC PRELIMINARY; PRT; 3011 AA.
AC Q9D7D8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
```

RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RL with hepatocellular carcinoma: the 'progression score' revisited.";
RN Hepatol. Res. 20:161-171(2001).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RT Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RA "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RN J. Gen. Virol. 72:2805-2809(1991).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RN J. Gen. Virol. 74:661-668(1993).
RL EMBL; AB049099; BAB18812.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR SMR; Q9DTE3; 1030-1658, 2009-2171, 2421-2950.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
GO; GO:0003723; F:RNA binding; IEA.
GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F:serine-type peptidase activity; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
GO; GO:0006350; P:transcription; IEA.
GO; GO:0019079; P:viral genome replication; IEA.
GO; GO:0019087; P:viral transformation; IEA.
GO; GO:0019040; DEAD.
GO; GO:0011545; DEAD/DEAH N.
GO; GO:0002521; HCV core.
GO; GO:0002519; HCV env.
GO; GO:0002531; HCV NS1.
GO; GO:000745; HCV NS4a.
GO; GO:0001490; HCV NS4b.
GO; GO:0002868; HCV NS5a.
GO; GO:0001650; Helicase C.
GO; GO:0002518; Pept U39_HCV NS2.
GO; GO:0004109; Peptidase S29.
GO; GO:0007095; RNA pol_PS.
GO; GO:0007094; RNA pol_Psivir.
GO; GO:001543; HCV capsid; 1.
GO; GO:001542; HCV core; 1.
GO; GO:001539; HCV env; 1.
GO; GO:001560; HCV NS1; 1.
GO; GO:001538; HCV NS2; 1.
GO; GO:002907; HCV NS3; 1.
GO; GO:001006; HCV NS4a; 1.
GO; GO:001001; HCV NS4b; 1.
GO; GO:001506; HCV NS5a; 1.
GO; GO:002271; Helicase C; 1.
GO; GO:000998; RDRP_3; 1.
GO; GO:000487; DEXDC; 1.
GO Polyprotein.
SQ SEQUENCE 3011 AA; 326853 MW; 068D3BDF753F184D CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred.No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
|||||
DB 1665 GGVLAALAAAYCL 1676

RESULT 275
Q9DTE3_9HEPC PRELIMINARY; PRT; 3011 AA.
ID Q9DTE3;
AC Q9DTE3;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RL with hepatocellular carcinoma: the 'progression score' revisited.";
RN Hepatol. Res. 20:161-171(2001).
[2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RL single patient show sequence heterogeneity.";
RN J. Gen. Virol. 72:2805-2809(1991).
[3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RN J. Gen. Virol. 74:661-668(1993).
RL EMBL; AB049094; BAB18807.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0250; PQ0250.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0253; PQ0253.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0255; PQ0255.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; 1CW.
DR SMR; Q9DTE3; 1030-1658, 2009-2171, 2421-2950.
GO; GO:0019028; C:viral capsid; IEA.
GO; GO:0019031; C:viral envelope; IEA.
GO; GO:0005524; F:ATP binding; IEA.
GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
GO; GO:0003723; F:RNA binding; IEA.
GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
GO; GO:0008236; F:serine-type peptidase activity; IEA.
GO; GO:0005198; F:structural molecule activity; IEA.
GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
GO; GO:0006350; P:transcription; IEA.
GO; GO:0019079; P:viral genome replication; IEA.
GO; GO:0019087; P:viral transformation; IEA.
GO; GO:0019040; DEAD.
GO; GO:0011545; DEAD/DEAH N.
GO; GO:0002521; HCV capsid.
GO; GO:0002519; HCV env.
GO; GO:0002531; HCV NS1.
GO; GO:000745; HCV NS4a.
GO; GO:0001490; HCV NS4b.
GO; GO:0002868; HCV NS5a.
GO; GO:0001650; Helicase C.
GO; GO:0002518; Pept U39_HCV NS2.
GO; GO:0004109; Peptidase S29.
GO; GO:0007095; RNA pol_PS.
GO; GO:0007094; RNA pol_Psivir.
GO; GO:001543; HCV capsid; 1.
GO; GO:001542; HCV core; 1.
GO; GO:001539; HCV env; 1.
GO; GO:001560; HCV NS1; 1.
GO; GO:001538; HCV NS2; 1.
GO; GO:002907; HCV NS3; 1.
GO; GO:001006; HCV NS4a; 1.
GO; GO:001001; HCV NS4b; 1.
GO; GO:001506; HCV NS5a; 1.
GO; GO:002271; Helicase C; 1.
GO; GO:000998; RDRP_3; 1.
GO; GO:000487; DEXDC; 1.
GO Polyprotein.
SQ SEQUENCE 3011 AA; 326853 MW; 068D3BDF753F184D CRC64;

```
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3011 AA; 326819 MW; 70434051PF91303B CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3011;
Best Local Similarity 100.0%; Pred. NO. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1665 GGVLAALAAAYCL 1676

RESULT 276
Q9WIK7_9HEPC
ID Q9WIK7_9HEPC PRELIMINARY; PRT; 3012 AA.
AC Q9WIK7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Rappeter K.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Rappeter K.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RX Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
RX Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "CDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=835694;
RX Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RX Dubeiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
RN [5]
RX EMBL; AJ132996; CAB41950.1; -; Genomic_RNA.
```

```
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR PIR; PQ329; PQ329.
DR HSP; Q8JYK1; 1CWX.
DR SRR; Q9WIK7; 1031-1659, 2010-2172, 2422-2951.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; F:ATP binding; IEA.
DR GO; GO:0005524; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0008026; F:RNA binding; IEA.
DR GO; GO:0003723; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3012 AA; 326827 MW; 7E59DB735B5C5115 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3012;
Best Local Similarity 100.0%; Pred. NO. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
Db 1666 GGVLAALAAAYCL 1677

RESULT 277
Q6J6P5_9HEPC
ID Q6J6P5_9HEPC PRELIMINARY; PRT; 3013 AA.
AC Q6J6P5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyprotein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Yao X., Guo J., Zheng C.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
```

```
RN NUCLEOTIDE SEQUENCE.
RA Li X., Zheng C.;
RL EMBL: AY587016; AAT40682.1; -; mRNA.
DR HSP: P26663; 1JXP.
DR SMR: Q63695; 1029-1657, 2008-2170, 2420-2949.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0019031; C: viral envelope; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F: RNA binding; IEA.
DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR GO: GO:0016740; F: transferase activity; IEA.
DR GO: GO:0006508; P: proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P: transcription; IEA.
DR GO: GO:0019079; F: viral genome replication; IEA.
DR GO: GO:0019087; P: viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV env.
DR InterPro: IPR002519; HCV NS1.
DR InterPro: IPR002531; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR007095; RNA pol_D5_PS.
DR InterPro: IPR007094; RNA pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01549; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3013 AA; 327183 MW; C71FA2B7C5257F3D CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAALAAAYCL 29
      |||||
DB 1664 GGVLAALAAAYCL 1675

RESULT 278
Q9J3H4_9HEPC
ID Q9J3H4_9HEPC PRELIMINARY; PRT; 3013 AA.
AC Q9J3H4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
```

```
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=MD18;
RA Nagayama K., Kurosaki M., Enomoto N., Miyaseaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=1314471;
RA Kato N., Ootsuyama Y., Tanaka T., Nakagawa M., Nakazawa T.,
RA Muraishi K., Ohkoshi S., Hijikata M., Shimotohno K.;
RT "Marked sequence diversity in the putative envelope proteins of
RT hepatitis C viruses.";
RL Virus Res. 22:107-123(1992).
DR EMBL; AF207759; AAF65949.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR PIR; S24097; S24097.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q8J3H4; 1029-1657, 2008-2170, 2423-2952.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0019031; C: viral envelope; IEA.
DR GO: GO:0005524; F: ATP binding; IEA.
DR GO: GO:0008026; F: ATP-dependent helicase activity; IEA.
DR GO: GO:0003723; F: RNA binding; IEA.
DR GO: GO:0003968; F: RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F: serine-type peptidase activity; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR GO: GO:0006350; P: proteolysis and peptidolysis; IEA.
DR GO: GO:0019079; P: viral genome replication; IEA.
DR GO: GO:0019087; P: viral transformation; IEA.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR011545; DEAD/DEAH N.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRp.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR007095; RNA pol_DS_PS.
DR InterPro: IPR007094; RNA pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3013 AA; 326960 MW; E4F80F8EA0E5C1E5 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred. No. 0.032;
```

```
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1664 GGVLAALAAAYCL 1675

RESULT 279
Q9QIX9_9HEPC
ID Q9QIX9_9HEPC PRELIMINARY; PRT; 3013 AA.
AC Q9QIX9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptin.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD6-2;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9204457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
J. Gen. Virol. 74:661-668(1993).
RL EMBL; AF165056; AAD56191.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0253; PQ0253.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0255; PQ0255.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; O8JYSL; 1CW.
DR SMR; Q9QIX9; 1032-1660, 2011-2173, 2423-2952.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DRAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
```

```
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39_HCV NS2.
DR InterPro; IPR004109; Peptidase S23.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01539; HCV_core; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW POLYPEPTIN.
SQ SEQUENCE 3013 AA; 326923 MW; 98D0BDE208A9B90E CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred. NO. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 18 GGVLAALAAAYCL 29
Db 1667 GGVLAALAAAYCL 1678

RESULT 280
Q9QIY0_9HEPC
ID Q9QIY0_9HEPC PRELIMINARY; PRT; 3013 AA.
AC Q9QIY0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptin.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD6-1;
RX MEDLINE=20013325; PubMed=10544098; DOI=10.1006/viro.1999.9924;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.Y., Miyasaka Y.,
Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=9204457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M.,
Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
Dusheiko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
J. Gen. Virol. 74:661-668(1993).
RL EMBL; AF165055; AAD56190.1; -; Genomic_RNA.
```

DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSSP; Q8JYS1; ICWX.
DR SMR; Q8JYS1; 1032-1660, 2011-2173, 2423-2952.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS4a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3013 AA; 326890 MW; 762E2D4B6B607B8C CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 18 GGVLAALAAAYCL 29
|||||
Db 1667 GGVLAALAAAYCL 1678

RESULT 281
Q9QNC0_9HEPC
ID Q9QNC0_9HEPC PRELIMINARY; PRT; 3013 AA.
AC Q9QNC0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=99428762; PubMed=10497111; DOI=10.1006/viro.1999.9879;
RA Zhang J., Yamada O., Ito T., Akiyama M., Hashimoto Y., Yoshida H.,
RA Makino R., Masago A., Uemura H., Araki H.;
RT "A single nucleotide insertion in the 5'-untranslated region of
RT hepatitis C virus leads to enhanced cap-independent translation.";
RL Virology 261:263-270(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagaaki M., Orita T., Hasegawa M.,
RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a
RT single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93224886; PubMed=8385694;
RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
RA Dushenko G., Saeed A.A., Holmes E.C.;
RT "Sequence variability in the 5' non-coding region of hepatitis C
RT virus: identification of a new virus type and restrictions on sequence
RT diversity.";
RL J. Gen. Virol. 74:661-668(1993).
DR EMBL; AB016785; BAA83719.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR SMR; Q9QNC0; 1030-1658, 2009-2171, 2423-2952.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS4a.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01538; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3013 AA; 327096 MW; 552D2C4EAD326DC CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      18 GGVLAALAAAYCL 29
Db      1665 GGVLAALAAAYCL 1676

RESULT 282
Q6GYR8_9HEPC PRELIMINARY; PRT; 3014 AA.
AC Q6GYR8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-RF1.2K/1b; DOI=10.1099/vir.0.79984-0;
RX PubMed=15218169;
RA Kalinina O., Norder H., Magnus L.O.;
RT "Full-length open reading frame of a recombinant hepatitis C virus
RT strain from St Petersburg: proposed mechanism for its formation.";
RL J. Gen. Virol. 85:1853-1857(2004).
DR EMBL; AY587845; AAT00644.1; -; Genomic_RNA.
DR SMR; Q6GYR8; 1033-1661; 2424-2953.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3014 AA; 328025 MW; 47080FE1AA55EBBP CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3014;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      18 GGVLAALAAAYCL 29
Db      1668 GGVLAALAAAYCL 1679

RESULT 283
Q86614_9HEPC PRELIMINARY; PRT; 3014 AA.
AC Q86614;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein protein.
OS Name-polyprotein;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93286423; PubMed=8393788;
RA Hayashi N., Higashi H., Kamataka K., Sugimoto H., Esumi M.,
RA Komatsu K., Hayashi K., Sugitani M., Suzuki K., Tadao O. et. al.;
RT "Molecular cloning and heterogeneity of the human hepatitis C virus
RT (HCV) genome.";
RL J. Hepatol. 17:S94-S107(1993).
DR EMBL; S62220; AAB27127.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; IDXP.
DR SMR; Q86614; 1029-1657; 2424-2953.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
SQ SEQUENCE 3014 AA; 326692 MW; 5C5E677DEFEB377A CRC64;
```



```
Query Match      10.2%; Score 12; DB 2; Length 3014;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      18 GGVLAAALAAAYCL 29
Db      1664 GGVLAAALAAAYCL 1675

RESULT 284
O9DTE0_9HEPC PRELIMINARY; PRT; 3014 AA.
AC O9DTE0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Serum;
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
RA Mishiro S.;
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients
RT with hepatocellular carcinoma: the 'progression score' revisited.";
RL Hepatol. Res. 20:161-171(2001).
DR EMBL; AB049097; BAB19810.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSSP; Q81755; 1DXP.
DR SMR; O9DTE0; 1029-1657, 2424-2953.
DR MEROPS; S29.002; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
PFam; PF01543; HCV capsid; 1.
PFam; PF01542; HCV core; 1.
PFam; PF01539; HCV env; 1.
PFam; PF01560; HCV NS1; 1.
PFam; PF01538; HCV NS2; 1.
PFam; PF02907; HCV NS3; 1.
PFam; PF01006; HCV NS4a; 1.
PFam; PF01003; HCV NS4b; 1.
PFam; PF01506; HCV NS5a; 1.
PFam; PF00271; Helicase_C; 1.

DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3014 AA; 327498 MW; ED9BF4F94BDB6287 CRC64;

Query Match      10.2%; Score 12; DB 2; Length 3014;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      18 GGVLAAALAAAYCL 29
Db      1664 GGVLAAALAAAYCL 1675

RESULT 285
O9WPH5_9HEPC PRELIMINARY; PRT; 3015 AA.
AC O9WPH5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-N;
RX MEDLINE=99315771; PubMed=10385673; DOI=10.1002/hep.510300137;
RA Beard M.R., Abell G., Honda M., Carroll A., Gartland M., Clarke B.,
RA Suzuki K., Lanford R., Sangar D.V., Lemon S.M.;
RT "An infectious molecular clone of a Japanese genotype 1b hepatitis C
RT virus ";
RL Hepatology 30:316-324(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-N;
RA Beard M.R., Abell G., Honda M., Carroll A., Gartland M., Clarke B.,
RA Suzuki K., Lanford R., Sangar D.V., Lemon S.M.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF139594; AAD44718.2; -; Genomic_RNA.
DR HSSP; Q8JYS1; 1CWX.
DR SMR; O9WPH5; 1030-1658, 2425-2954.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; P:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol_DS_PS.
DR InterPro; IPR007094; RNA pol_PSVir.
PFam; PF01543; HCV capsid; 1.
PFam; PF01542; HCV core; 1.
```

DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXdc; 1.
KW Polyprotein.
SQ SEQUENCE 3015 AA; 326625 MW; A86AE71196578EE3 CRC64;

Query Match 10.2%; Score 12; DB 2; Length 3015;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1665 GGVLAALAAAYCL 1676

RESULT 286

O92529_9HEPC
ID O92529_9HEPC PRELIMINARY; PRT; 3019 AA.
AC O92529;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TB580;
RX MEDLINE=98378034; PubMed=9714232;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuruta F., Miyakawa Y., Mayumi M.;
RT "The entire nucleotide sequences of three hepatitis C virus isolates in genetic groups 7-9 and comparison with those in the other eight genetic groups.";
RL J. Gen. Virol. 79:1847-1857 (1998).
DR EMBL; D84262; BAA32664.1; -; Genomic_RNA.
DR HSP; Q8JY81; ICW.
DR SMR; O92529; 1034-1662, 2429-2994.
DR MEROPS; C18.001; -.
DR CO; GO:0019028; C:viral capsid; IEA.
DR CO; GO:0019031; C:viral envelope; IEA.
DR CO; GO:0005524; F:ATP binding; IEA.
DR CO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR CO; GO:0003723; F:RNA binding; IEA.
DR CO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR CO; GO:0008236; F:serine-type peptidase activity; IEA.
DR CO; GO:0005198; F:structural molecule activity; IEA.
DR CO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR CO; GO:0006350; P:transcription; IEA.
DR CO; GO:0019079; P:viral genome replication; IEA.
DR CO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.

DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXdc; 1.
KW Polyprotein.
SQ SEQUENCE 3019 AA; 328232 MW; E26750E07BCBC310 CRC64;
Query Match 10.2%; Score 12; DB 2; Length 3019;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALAAAYCL 29
|||||
Db 1669 GGVLAALAAAYCL 1680

RESULT 287
O91AU0_9HEPC
ID O91AU0_9HEPC PRELIMINARY; PRT; 3010 AA.
AC O91AU0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21440119; PubMed=11556407; DOI=10.1023/A:1011143731677;
RA Lim S.P., Khu Y.L., Hong W.J., Tay A.E., Ting A.E., Lim S.G., Tan Y.H.;
RT "Identification and molecular characterization of the complete genome of a Singapore isolate of hepatitis C virus: sequence comparison with other strains and phylogenetic analysis.";
RL Virus Genes 23:89-95 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92044457; PubMed=1658209;
RA Oshima M., Teuchiya M., Yagasaki M., Orita T., Hasegawa M., Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
RT "cDNA clones of Japanese hepatitis C virus genomes derived from a single patient show sequence heterogeneity.";
RL J. Gen. Virol. 72:2805-2809 (1991).
DR EMBL; AF356827; AAL00900.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; Q81755; IDXP.
DR SMR; O91AU0; 1029-1657, 2008-2170, 2420-2949.
DR CO; GO:0019028; C:viral capsid; IEA.
DR CO; GO:0019031; C:viral envelope; IEA.
DR CO; GO:0005524; F:ATP binding; IEA.
DR CO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR CO; GO:0003723; F:RNA binding; IEA.
DR CO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR CO; GO:0008236; F:serine-type peptidase activity; IEA.
DR CO; GO:0005198; F:structural molecule activity; IEA.
DR CO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR CO; GO:0006350; P:transcription; IEA.
DR CO; GO:0019079; P:viral genome replication; IEA.
DR CO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.

```

DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3010 AA; 326798 MW; 3D89304314F9F795 CRC64;

Query Match 9.3%; Score 11; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALAAAYCL 29
Db 1665 GVLAALAAAYCL 1675

RESULT 288
Q68798_9HEPC
ID Q68798_9HEPC PRELIMINARY; PRT; 3022 AA.
AC Q68798.
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96226020; PubMed=8627233;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tsuda F.,
RA Lesmana L.A., Miyakawa Y., Mayumi M.;
RT "Hepatitis C virus variants from Jakarta, Indonesia classifiable into
RT novel genotypes in the second (2e and 2f), tenth (10a) and eleventh
RT (11a) genetic groups.";
RL J. Gen. Virol. 77:293-301(1996).
DR EMBL; D63822; BAA09891.1; -; Genomic_RNA.
DR HSSP; F26663; 1QVU.
DR SMR; Q68798; 1033-1661, 2432-2996.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:serine-type RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.

```

```

DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3022 AA; 328694 MW; EE6A21538FEA26E1 CRC64;

Query Match 9.3%; Score 11; DB 2; Length 3022;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 YQYDEMEECs 66
Db 1706 YQYDEMEECs 1716

RESULT 289
Q988U6_RHILO
ID Q988U6_RHILO PRELIMINARY; PRT; 434 AA.
AC Q988U6.
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE M116588 protein.
GN OrderedLocusNames=m116588;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=21089930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000012; BAB52851.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 434 AA; 43750 MW; 5FB9B6D99C26BD39 CRC64;

Query Match 8.5%; Score 10; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLAALA 25
Db 63 LLGGVLAALA 72

```

```

RESULT 290
Q9HRP0_HALSA
ID Q9HRP0_HALSA PRELIMINARY; PRT; 280 AA.
AC Q9HRP0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 4-hydroxybenzoate octaprenyltransferase.
GN Name=rhoA; OrderedLocusNames=VNG0610G;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Saliga N.S., Thoreson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Iisenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005009; AAG19118.1; -; Genomic_DNA.
DR PIR; B84219; B84219.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR000537; UbiA_prenyltrans.
DR Pfam; PF01040; UbiA; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 280 AA; 7635 MW; B6DE58FD22235A03 CRC64;

Query Match 7.6%; Score 9; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAA 26
Db 161 GGVLAAALAA 169

RESULT 291
Q81242_9HEPC
ID Q81242_9HEPC PRELIMINARY; PRT; 484 AA.
AC Q81242;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Precursor protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BE95;
RX MEDLINE=95023999; PubMed=7524083;
RA Stuyver L., van Arnhem W., Wyseur A., Hernandez F., Delaporte E.,
RA Maertens G.;
RT "Classification of hepatitis C viruses based on phylogenetic analysis
RT of the envelope 1 and nonstructural 5B regions and identification of
RT five additional subtypes.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:10134-10138(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BE95;
RX MEDLINE=94338342; PubMed=7520237;

Query Match 7.6%; Score 9; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAA 26
Db 161 GGVLAAALAA 169

RESULT 292
Q5Z1V4_NOCFA
ID Q5Z1V4_NOCFA PRELIMINARY; PRT; 545 AA.
AC Q5Z1V4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative acyl-CoA synthetase.
GN OrderedLocusNames=nfa7120;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152; DOI=10.1073/pnas.0406410101;
RX PubMed=15466710; Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RX Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RX Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
CC -1- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
CC family.
DR EMBL; AP006618; BAD55557.1; -; Genomic DNA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000873; AMP-bind.
DR Pfam; PF00501; AMP-binding; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00455; AMP_BINDING.
KW Complete proteome.
SQ SEQUENCE 545 AA; 58381 MW; 41817428C63A48AB CRC64;

Query Match 7.6%; Score 9; DB 2; Length 545;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAALAA 26
Db 161 GGVLAAALAA 169

```

```

Db      280 GGVLAALAA 288
RESULT 293
Q6AH48 LEIXX
ID      Q6AH48_LEIXX PRELIMINARY;      PRT; 1235 AA.
AC      Q6AH48;
DT      25-OCT-2004 (TREMBLrel. 28, Created)
DT      25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DE      Proline dehydrogenase.
GN      Name=poaA; OrderedLocusNames=Lxx02530;
OS      Leifsonia xyl (subsp. xyl).
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Micrococcineae; Microbacteriaceae; Leifsonia.
OX      NCBI_TaxID=59736;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CTCB07;
RX      PubMed=15305603;
RA      Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA      Katajama J.P., Truffi D., do Amaral A.M., Harakava R.,
RA      de Oliveira J.C.P., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA      Takita M.A., da Silva A.C.R., Furian L.R., Carraro D.M., Camarotte G.,
RA      Almeida N.F. Jr., Carrier H., Coutinho L.B., El-Dorri H.A.,
RA      Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA      Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,
RA      Lemos M.V.P., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
RA      Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA      Tsai S.M., Zanca A.S., Simpson A.J.G., Brumley S.M., Secubal J.C.;
RT      "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
RT      xyl subsp. xyl.";
RL      Mol. Plant Microbe Interact. 17:827-836(2004).
DR      EMBL; AE016822; AAT88297.1; -; Genomic DNA.
DR      GO; GO:0016491; P:oxidoreductase activity; IEA.
DR      GO; GO:0004657; P:proline dehydrogenase activity; IEA.
DR      GO; GO:0006537; P:glutamate biosynthesis; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      GO; GO:0006562; P:proline catabolism; IEA.
DR      InterPro; IPR002086; Aldehyd dehydrog.
DR      InterPro; IPR002872; Prolin dehydrog.
DR      Pfam; PF00171; Aldehyd; 1.
DR      Pfam; PF01619; Pro dh; 1.
DR      PROSITE; PS00070; ALDEHYDE DEHYDR CYS; UNKNOWN 1.
DR      PROSITE; PS00687; ALDEHYDE DEHYDR GLU; UNKNOWN 1.
KW      Complete proteome.
SQ      SEQUENCE 1235 AA; 132947 MW; 3C1E2F1F4089A3FB CRC64;

Query Match      7.6%; Score 9; DB 2; Length 1235;
Best Local Similarity 100.0%; Pred.No.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAA 26
Db      711 GGVLAALAA 719
|||||||
RESULT 294
O39928_9HEPC
ID      O39928_9HEPC PRELIMINARY;      PRT; 3014 AA.
AC      O39928;
DT      01-JAN-1998 (TREMBLrel. 05, Created)
DT      01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Polyprotein.
OS      Hepatitis C virus type 5a.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
OX      NCBI_TaxID=31654;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=EUH1480;
RX      MEDLINE=97366593; PubMed=9223423; DOI=10.1006/bbrc.1997.6902;

Query Match      7.6%; Score 9; DB 2; Length 1235;
Best Local Similarity 100.0%; Pred.No.14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAALAA 26
Db      711 GGVLAALAA 719
|||||||
RESULT 295
O91936_9HEPC
ID      O91936_9HEPC PRELIMINARY;      PRT; 3014 AA.
AC      O91936;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Polyprotein.
OS      Hepatitis C virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
OX      NCBI_TaxID=11103;

```

```

RA      Chamberlain R.W., Adams N.J., Taylor L.A., Simmonds P., Elliot R.M.;
RT      "The complete coding sequence of hepatitis C virus genotype 5a, the
RT      predominant genotype in South Africa.";
RL      Biochem. Biophys. Res. Commun. 236:44-49(1997).
DR      EMBL; Y13184; CAA73640.1; -; Genomic_RNA.
DR      FIR; JC5620; JC5620.
DR      HSSP; O8JYS1; 1CWX.
DR      SMR; O39928; 1030-1658, 2424-2989.
DR      MEROPS; S29.001; -.
DR      GO; GO:0019028; C:viral capsid; IEA.
DR      GO; GO:0019031; C:viral envelope; IEA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR      GO; GO:0003723; F:RNA binding; IEA.
DR      GO; GO:0003168; F:RNA-directed RNA polymerase activity; IEA.
DR      GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR      GO; GO:0005198; F:structural molecule activity; IEA.
DR      GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR      GO; GO:0006350; P:transcription; IEA.
DR      GO; GO:0019079; P:viral genome replication; IEA.
DR      GO; GO:0019087; P:viral transformation; IEA.
DR      InterPro; IPR001410; DEAD.
DR      InterPro; IPR011545; DEAD/DEAH_N.
DR      InterPro; IPR002522; HCV capsid.
DR      InterPro; IPR002521; HCV core.
DR      InterPro; IPR002519; HCV env.
DR      InterPro; IPR002531; HCV NS1.
DR      InterPro; IPR000745; HCV NS4a.
DR      InterPro; IPR001490; HCV NS4b.
DR      InterPro; IPR002868; HCV NS5a.
DR      InterPro; IPR002166; HCV RdRP.
DR      InterPro; IPR001650; Helicase_C.
DR      InterPro; IPR004109; Peptidase_S29.
DR      InterPro; IPR002518; Pept_U39_HCV_NS2.
DR      InterPro; IPR007095; RNA_pol_DS_PS.
DR      InterPro; IPR007094; RNA_pol_PSVir.
DR      Pfam; PF01543; HCV capsid; 1.
DR      Pfam; PF01542; HCV core; 1.
DR      Pfam; PF01539; HCV env; 1.
DR      Pfam; PF01538; HCV NS2; 1.
DR      Pfam; PF01560; HCV NS1; 1.
DR      Pfam; PF02907; HCV NS3; 1.
DR      Pfam; PF01006; HCV NS4a; 1.
DR      Pfam; PF01001; HCV NS4b; 1.
DR      Pfam; PF01506; HCV NS5a; 1.
DR      Pfam; PF00271; Helicase_C; 1.
DR      Pfam; PF00998; RdRP_3; 1.
DR      SMART; SM00487; DEXDc; 1.
KW      Polyprotein.
SQ      SEQUENCE 3014 AA; 327668 MW; 84934E2B77A1964B CRC64;

Query Match      7.6%; Score 9; DB 2; Length 3014;
Best Local Similarity 100.0%; Pred.No.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACMSADLEV 9
Db      1648 ACMSADLEV 1656
|||||||
RESULT 295
O91936_9HEPC
ID      O91936_9HEPC PRELIMINARY;      PRT; 3014 AA.
AC      O91936;
DT      01-NOV-1998 (TREMBLrel. 08, Created)
DT      01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT      01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE      Polyprotein.
OS      Hepatitis C virus.
OC      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC      Hepacivirus.
OX      NCBI_TaxID=11103;

```

RN NUCLEOTIDE SEQUENCE.
RC STRAIN=8A13,
RX MEDLINE=99033006; PubMed=9806059;
RA Bukh J., Apper C.L., Engle R., Govindarajan S., Hegerich P.A.,
RA Tellier R., Wong D.C., Elkins R., Kew M.C.;
RT "Experimental infection of chimpanzees with hepatitis C virus of
RT genotype 5a: genetic analysis of the virus and generation of a
RT standardized challenge pool.";
RL J. Infect. Dis. 178:1193-1197(1998).
DR EMBL; AF064490; AAC61696.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CW.
DR SMR; O91936; 1030-1658, 2424-2953.
DR MEROPS; C18.001; -.
DR MEROPS; S29.001; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003688; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR02522; HCV capsid.
DR InterPro; IPR02521; HCV_core.
DR InterPro; IPR02519; HCV_env.
DR InterPro; IPR02531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR02868; HCV_NS5a.
DR InterPro; IPR02166; HCV_RdRp.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RdRp_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3014 AA; 327115 MW; 4D198683058C13D6 CRC64;

Query Match 7.6%; Score 9; DB 2; Length 3014;
Best Local Similarity 100.0%; Pred.No.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACMSADLEV 9
Db 1648 ACMSADLEV 1656

RESULT 296
Q4H710_9DEIO
ID Q4H710_9DEIO PRELIMINARY; PRT; 117 AA.
AC Q4H710;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein precursor.
GN ORFNames=DgeODRAFT_0345;

OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000009; EAL82245.1; -; Genomic_DNA.
KW Hypothetical protein; Signal.
FT SIGNAL 1 28 Potential.
SQ SEQUENCE 117 AA; 11930 MW; B0C01313785B5D23 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 117;
Best Local Similarity 100.0%; Pred.No.20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALA 25
Db 10 GGVLAALA 17

RESULT 297
Q68234_9HEPC
ID Q68234_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68234;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=1a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14270; AAC53959.1; -; Genomic_RNA.
DR HSP; P27958; IHEI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON TER 1
FT NON TER 138 138
SQ SEQUENCE 138 AA; 15365 MW; 117C12B8D0D29E19 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 138;
Best Local Similarity 100.0%; Pred.No.22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
DB 22 CMSADLEV 29

RESULT 298

QAP348 USTWA PRELIMINARY; PRT; 186 AA.

AC QAP348;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=U05465.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=521;
RA Birren B., Nuebaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blichsteyn B., Bloom I., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer G.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G.,
RA Gierke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huseby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kamat A., Kamyselis M., Karlsson E.,
RA Kellis C., Kieu A., Kisner P., Kodira C., Kulbokas B., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalov A., Mihova T., Mikkelsen T., Mienga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA O'Neill K., O'donnell P., Okaawa O., O'leary S., Omotosho B.,
RA Purcell S., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Rettar R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tensing P., Tesfaye S., Theodore J., Thoulteang Y., Topham K.,
RA Towey S., Tsamila T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACF01000197; EAK96398.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 186 AA; 20439 MW; 52D73F8405875BB6 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLAAL 24

DB 121 LGGVLAAL 128

RESULT 299

Q4UXM7_XANCP PRELIMINARY; PRT; 215 AA.

AC Q4UXM7;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=XC_1126;
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
RE EMBL; CP000050; AAY48196.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 215 AA; 22796 MW; 08F5EBB960132F4E CRC64;

Query Match 6.8%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAAYC 28

DB 47 LAALAAYC 54

RESULT 300

Q8P6D8_XANCP PRELIMINARY; PRT; 215 AA.

AC Q8P6D8;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein XCC3032.
GN OrderedLocustNames=XCC3032;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Bertolini M.C.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lenos E.G.M., Lenos M.V.F.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."

```

RL Nature 417:459-463(2002).
DR EMBL; AB012416; AAM42303.1; -; Genomic_DNA.
KW Complete proteome, Hypothetical protein.
SQ SEQUENCE 215 AA; 22766 MW; F2AB5BBD013344F CRC64;

Query Match 6.8%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LALAAAYC 28
Db 47 LALAAAYC 54

RESULT 301
ID QANV17_9DELTA PRELIMINARY; PRT; 220 AA.
AC QANV17;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Similar to butyrate response factor 2; EGF-response factor 2; zinc
DE finger protein, C3H type, 36-like 2 precursor.
GN ORFNames=AdelDRAFT_2501;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cyto bacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-JGFP);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Terani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAMD0100010; EAL79707.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL 1 20 Potential.
SQ SEQUENCE 220 AA; 20579 MW; 6C56B0A2AD748A0B CRC64;

Query Match 6.8%; Score 8; DB 2; Length 220;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALAA 26
Db 66 GVLAALAA 73

RESULT 302
ID Q93HX1 MAGMG PRELIMINARY; PRT; 279 AA.
AC Q93HX1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Iron-sulfur protein.
GN Name=naph;
OS Magnetospirillum magnetotacticum (Aquaspirillum magnetotacticum).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;

```

```

OC Rhodospirillaceae; Magnetospirillum.
OX NCBI_TaxID=188;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MS-1;
RA Taoka A., Fukumori Y.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Involved in electron transfer (by similarity).
DR EMBL; AB055444; BAB59024.1; -; Genomic_DNA.
DR HSSP; P55907; IXER.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR011450; 4Fe4S_Fe_S_bd.
DR InterPro; IPR011886; Naph_.
DR Pfam; PF00037; Fer4; 1.
DR TIGRFAMs; TIGR02163; naph; 1.
DR PROSITE; PS00198; 4Fe4S-FERREDOXIN; 1.
KW 4Fe-4S; Electron transport; Iron; Iron-sulfur; Metal-binding;
KW Transport.
SQ SEQUENCE 279 AA; 29610 MW; 5483C05BECBAC4D1 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 23
Db 130 LLGGVLA 137

RESULT 303
ID Q8PWS7 METMA PRELIMINARY; PRT; 296 AA.
AC Q8PWS7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Methionine aminopeptidase (EC 3.4.11.18).
GN OrderedLocusNames=MM1499;
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Bosmecke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL; AB013384; AAM31195.1; -; Genomic_DNA.
DR HSSP; P56218; 1XGS.
DR MEROPS; M24.002; -.
DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001714; Pept_M24 MAP2.
DR InterPro; IPR002468; Pept_M24A MAP2.
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF00557; Peptidase_M24; 1.
DR PRINTS; PR00599; MAPEPTIDASE.
DR TIGRFAMs; TIGR00501; met_pdaseII; 1.
DR PROSITE; PS01202; MAP_2; UNKNOWN 1.
KW Aminopeptidase; Cobalt; Complete proteome; Hydrolase; Protease.
SQ SEQUENCE 296 AA; 32166 MW; 0039FD7031600ECO CRC64;

```



```
Query Match          6.8%; Score 8; DB 2; Length 296;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 IELGGKPA 47
    |||||
Db 51 IELGGKPA 58

RESULT 304
Q5YXT2_NOCFA PRELIMINARY; PRT; 305 AA.
AC Q5YXT2_
DT 25-OCT-2004 (TREMELrel. 28, Created)
DT 25-OCT-2004 (TREMELrel. 28, Last sequence update)
DE Putative phenylacetic acid degradation protein.
GN OrderedLocusNames=nfa21630;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RL "The complete genomic sequence of Nocardia farcinica IFM 10152.";
DR EMBL; AP006618; BAD57009.1; -; Genomic_DNA.
DR InterPro; IPR007814; PaaA_PaaC.
DR InterPro; IPR011882; PA_CoA_Oxy3.
DR Pfam; PF05138; PaaA_PaaC; 1.
DR TIGRFAMs; TIGR02158; PA_CoA_Oxy3; 1.
KW Complete proteome.
SQ SEQUENCE 305 AA; 32628 MW; B9445949F5B599A4 CRC64;

Query Match          6.8%; Score 8; DB 2; Length 305;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYCL 29
    |||||
Db 44 AALAAAYCL 51

RESULT 305
Q9LIY9_STRCO PRELIMINARY; PRT; 320 AA.
AC Q9LIY9;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Putative integral membrane protein.
GN OrderedLocusNames=SCO2646; ORFNames=SC8E4A.16c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
```

```
RT "Complete genome sequence of the model actinomycete Streptomyces
RL coelicolor A3(2).";
DR EMBL; AL939113; CAB71821.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane, IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome.
SQ SEQUENCE 320 AA; 32034 MW; F52ABA71EBEPB716 CRC64;

Query Match          6.8%; Score 8; DB 2; Length 320;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAALAA 26
    |||||
Db 146 GVLAAALAA 153

RESULT 306
PHO36_YEAST STANDARD; PRT; 327 AA.
AC Q12442;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ADIPOR-like receptor PHO36 (Phosphate metabolism protein 36).
GN Name=PHO36; Synonyms=12H2; OrderedLocusNames=YOL002C; ORFNames=UND327;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97051599; PubMed=8896276;
RX DOI=10.1002/(SICI)1097-0061(199609)12:10B<1091::AID-YEA22>3.3.CO;2-9;
RA Sterky F., Holmberg A., Pettersson B., Uhlen M.;
RT "The sequence of a 30 kb fragment on the left arm of chromosome XV
RT from Saccharomyces cerevisiae reveals 15 open reading frames, five of
RT which correspond to previously identified genes.";
RL Yeast 12:1091-1095(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=97313270; PubMed=9169874;
RA Dujon B., Albermann K., Aldea M., Alexandraki D., Ansoerge W.,
RA Arino J., Benes V., Bohn C., Bolotin-Fukuhara M., Bordonne R.,
RA Boyer J., Camases A., Casamayor A., Casas C., Cheret G.,
RA Cziepluch C., Daigman-Fornier B., Dang D.V., de Haan M., Delius H.,
RA Durand P., Fairhead C., Feldmann H., Gaillon L., Galisson F.,
RA Gamo F.-J., Gancedo C., Goffeau A., Goulding S.E., Grivell L.A.,
RA Habbig B., Hand N.J., Hani J., Hattenhorst U., Hebling U.,
RA Hernandez Y., Herrero E., Heumann K., Hiesel R., Hilger F., Hofmann B.,
RA Hollenberg C.P., Hughes B., Jauniaux J.-C., Kalogeropoulos A.,
RA Katsoulou C., Kordes E., Lafuente M.J., Landt O., Louis E.J.,
RA Maarse A.C., Madania A., Mannhaupt G., Marck C., Martin R.P.,
RA Meves H.-W., Michaux G., Paces V., Parle-McDermott A.G., Pearson B.M.,
RA Perrin A., Pettersson B., Poch O., Pohl T.M., Poirey R.,
RA Portetelle D., Pujoil A., Purnelle B., Ramezani Rad M., Rechmann S.,
RA Schwager C., Schweizer M., Sor F., Sterky F., Tarassov I.A.,
RA Teodoru C., Tettelin H., Thierry A., Tobiasch E., Tzermias M.,
RA Uhlen M., Unsel M., Valens M., Vandenbol M., Vetter I., Vlcek C.,
RA Voet M., Volckaert G., Voss H., Wambutt R., Wedler H., Wiemann S.,
RA Winsor B., Wolfe K.H., Zollner A., Zumstein E., Kleine K.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XV.";
RL Nature 387:98-102(1997).
RN [3]
RP INDUCTION.
RX MEDLINE=98449944; PubMed=9774671;
RA Karpichev I.V., Small G.M.;
RT "Global regulatory functions of Oaf1p and Pip2p (Oaf2p), transcription
RT factors that regulate genes encoding peroxisomal proteins in
RT Saccharomyces cerevisiae.";
RL Mol. Cell. Biol. 18:6560-6570(1998).
```

```
RN
RP FUNCTION.
RA MEDLINE=22028123; PubMed=11916977; DOI=10.1074/jbc.M202045200;
RA Karpichev I.V., Cornivelli L., Small G.M.;
RT "Multiple regulatory roles of a novel Saccharomyces cerevisiae
RT protein, encoded by YOL002c, in lipid and phosphate metabolism."
RL J. Biol. Chem. 277:19609-19617(2002).
CC -1- FUNCTION: Probable receptor, which is involved in metabolic
CC pathways that regulate lipid metabolism such as fatty acid
CC oxidation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- INDUCTION: Regulated by OAP1 and PIP2. Highly expressed in
CC presence of glucose. Expressed at lower level in presence of
CC glycerol or glycerol and oleate. Highly expressed in the presence
CC of saturated fatty-acids such as myristate.
CC -1- SIMILARITY: Belongs to the ADIPOR family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
EMBL; U43491; AAC49478.1; -; Genomic DNA.
EMBL; Z74744; CAA93001.1; -; Genomic_DNA.
PIR; S61982; S61982.
Ensembl; YOL002C; Saccharomyces cerevisiae.
SGD; S000005362; IZH2.
GO; GO:0016020; C:membrane; TAS.
GO; GO:0006629; P:lipid metabolism; TAS.
GO; GO:0006882; P:zinc ion homeostasis; IMP.
DR InterPro; IPR004254; HlyIII-related.
DR Pfam; PF03006; HlyIII; 1.
DR Complete proteome; Fatty acid metabolism; Lipid metabolism; Receptor;
KW Transmembrane.
FT TOPO_DOM 1 88 Cytoplasmic (Potential).
FT TRANSMEM 89 109 Extracellular (Potential).
FT TOPO_DOM 110 120 Extracellular (Potential).
FT TRANSMEM 121 141 2 (Potential).
FT TOPO_DOM 142 163 Cytoplasmic (Potential).
FT TRANSMEM 164 184 3 (Potential).
FT TOPO_DOM 185 186 Extracellular (Potential).
FT TRANSMEM 187 207 4 (Potential).
FT TOPO_DOM 208 222 Cytoplasmic (Potential).
FT TRANSMEM 223 243 5 (Potential).
FT TOPO_DOM 244 252 Extracellular (Potential).
FT TRANSMEM 253 273 6 (Potential).
FT TOPO_DOM 274 286 Cytoplasmic (Potential).
FT TRANSMEM 287 307 7 (Potential).
FT TOPO_DOM 308 327 Extracellular (Potential).
SQ SEQUENCE 327 AA; 37437 MW; 08FE06D39AE2DD4 CRC64;

Query Match 6.8%; Score 8; DB 1; Length 327;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 WVLGGVVL 21
Db 257 WVLGGVVL 264

RESULT 307
Q57F22 BRUAB PRELIMINARY; PRT; 342 AA.
AC Q57F22;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical proline racemase.
GN OrderedLocuNames=BrAb1_0363;
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
```

```
OX NCBI_TaxID=235;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=9-941 / Biovar 1;
RX PubMed=15805518; DOI=10.1128/JB.187.8.2715-2726.2005;
RA Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuerner R.L.,
RA Qing Z., Li L.-L., Kapur V., Alt D.P., Olsen S.C.;
RT "Completion of the genome sequence of Brucella abortus and comparison
RT to the highly similar genomes of Brucella melitensis and Brucella
RT suis."
RL J. Bacteriol. 187:2715-2726(2005).
DR EMBL; AB017223; AAX73762.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 342 AA; 36941 MW; 1982D458A78A8E8A CRC64;

Query Match 6.8%; Score 8; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPAI 48
Db 297 ELGGKPAI 304

RESULT 308
Q8G2I3 BRUSU PRELIMINARY; PRT; 342 AA.
ID Q8G2I3 BRUSU PRELIMINARY;
AC Q8G2I3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Proline racemase, putative.
GN OrderedLocuNames=BR0337;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Unayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
DR EMBL; AB014291; AAN29286.1; -; Genomic_DNA.
DR SMR; Q8G2I3; 1-332.
DR TIGR; BR0337; -.
DR InterPro; IPR008794; Pro_racemase.
DR PANTHER; PTHR18835; Pro_racemase; 1.
DR Pfam; PF05544; Pro_racemase; 1.
DR PIRSF; PIRSF029792; Pro_racemase; 1.
KW Complete proteome.
SQ SEQUENCE 342 AA; 36941 MW; 1982D458A78A8E8A CRC64;

Query Match 6.8%; Score 8; DB 2; Length 342;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPAI 48
Db 297 ELGGKPAI 304

RESULT 309
Q8YFD6 BRUME PRELIMINARY; PRT; 342 AA.
ID Q8YFD6 BRUME PRELIMINARY;
```

AC Q8YFD6;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE PROLINE RACEMASE (EC 5.1.1.4).
 GN OrderedLocusNames=BMEI1566;
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX PubMed=1175688; DOI=10.1073/pnas.221573398;
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
 RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kypides N.C., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
 DR EMBL; AE009594; AAL52767.1; -; Genomic_DNA.
 DR PIR; AD3450; AD3450.
 DR PDB; 1TM0; X-ray; A/B=1-342.
 DR GO; GO:0018112; F:proline racemase activity; IEA.
 DR InterPro; IPR008794; Pro_racemase.
 DR PANTHER; PTHR18835; Pro_racemase; 1.
 DR Pfam; PF05544; Pro_racemase; 1.
 DR PIRSF; PIRSF029792; Pro_racemase; 1.
 KW Complete proteome.
 SQ SEQUENCE 342 AA; 36974 MW; FB6E5E9208EF98DC CRC64;

Query Match 6.8%; Score 8; DB 2; Length 342;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPAI 48
 |||||
 Db 297 ELGGKPAI 304

RESULT 310
 O31809_BACSU
 ID O31809_BACSU PRELIMINARY; PRT; 363 AA.
 AC O31809;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE YndE protein.
 GN Name=yndE; OrderedLocusNames=BSU17760;
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
 RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.-K., Codani J.-J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.-Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.-F., Itaya M.,
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
 RA Klein C., Kobayashi Y., Kottler P., Konharsky G., Krogh S.,
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
 RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,

RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
 RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
 RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,
 RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
 RA Rose M., Sadaie F., Sato T., Scanlan E., Schleicher S., Schroter R.,
 RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,
 RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
 RA Takemaru K., Takeuchi M., Takakoshi A., Tanaka T., Terpstra P.,
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
 RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
 RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
 RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,
 RA Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RT subtilis.";
 RL Nature 390:249-256 (1997).
 DR EMBL; Z99113; CAB13660.1; -; Genomic_DNA.
 DR PIR; E69889; E69889.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR GO; GO:0009847; P:spore germination; IEA.
 DR InterPro; IPR002293; AA/rei_permease1.
 DR InterPro; IPR004761; Spore_permease2.
 DR Pfam; PF03845; Spore_permease; 1.
 DR TIGRFAMs; TIGR00912; 2A0309; 1.
 KW Complete proteome; Transmembrane.
 SQ SEQUENCE 363 AA; 40470 MW; 1F69D5E3EF312EB5 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 363;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVLA 22
 |||||
 Db 46 VLLGGVLA 53

RESULT 311
 Q82CF6_STRAW
 ID Q82CF6_STRAW PRELIMINARY; PRT; 369 AA.
 AC Q82CF6;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=SAV5394;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Oonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbr820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";

```

RL  Nat. Biotechnol. 21:526-531(2003).
DR  EMBL; BA000030; BAC73106.1; -; Genomic_DNA.
DR  GO; GO:0016020; C:membrane; IEA.
DR  InterPro; IPR000620; DUF6_TM.
DR  Pfam; PF00892; DUF6; 2.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 369 AA; 36920 MW; 945DD9B79E75AAE2 CRC64;

Query Match          6.8%; Score 8; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  19 GVLAAALAA 26
    |||||
Db  149 GVLAAALAA 156

RESULT 312
Q7D4N4 MYCTU
ID  Q7D4N4 MYCTU PRELIMINARY; PRT; 399 AA.
AC  Q7D4N4;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  PPE family protein.
GN  OrderedLocusNames=MT4007;
OS  Mycobacterium tuberculosis.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC  Mycobacterium tuberculosis complex.
NCBI_TaxID=1773;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=CDC 1551 / Oshkosh;
RX  MEDLINE=22206494; PubMed=12218036;
RX  DOI=10.1128/JB.184.19.5479-5490.2002;
RA  Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA  Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA  Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA  Salzberg S.L., Delcher A., Weidman J.F., Khouri H.M.,
RA  Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA  Fraser C.M.;
RT  "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT  laboratory strains.";
RL  J. Bacteriol. 184:5479-5490(2002).
DR  EMBL; AR000516; AAK48374.1; -; Genomic_DNA.
DR  TIGR; MT4007; -.
DR  InterPro; IPR000030; Microbac_PPE.
DR  Pfam; PF00823; PPE; 1.
SQ  SEQUENCE 399 AA; 38837 MW; DA2B9D7A40C9C65 CRC64;

Query Match          6.8%; Score 8; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  18 GGVLAALA 25
    |||||
Db  160 GGVLAALA 167

RESULT 313
Q7TVE9 MYCBO
ID  Q7TVE9 MYCBO PRELIMINARY; PRT; 399 AA.
AC  Q7TVE9;
DT  01-OCT-2003 (TrEMBLrel. 25, Created)
DT  01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  PPE FAMILY PROTEIN.
GN  Name=PPB69; OrderedLocusNames=Mb3921c;
OS  Mycobacterium bovis.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC  Mycobacterium tuberculosis complex.

```

```

OX  NCBI_TaxID=1765;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=AF2122/97;
RX  MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA  Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA  Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA  Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA  Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT  "The complete genome sequence of Mycobacterium bovis.";
RL  Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR  EMBL; BX248347; CAP96107.1; -; Genomic_DNA.
DR  InterPro; IPR000030; Microbac_PPE.
DR  Pfam; PF00823; PPE; 1.
KW  Complete proteome.
SQ  SEQUENCE 399 AA; 38879 MW; 343E3034D41A8701 CRC64;

Query Match          6.8%; Score 8; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  18 GGVLAALA 25
    |||||
Db  160 GGVLAALA 167

RESULT 314
Q79F90 MYCTU
ID  Q79F90 MYCTU PRELIMINARY; PRT; 399 AA.
AC  Q79F90;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  PPE FAMILY PROTEIN.
GN  Name=PPB69; OrderedLocusNames=Rv3892c;
OS  Mycobacterium tuberculosis.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC  Mycobacterium tuberculosis complex.
NCBI_TaxID=1773;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RC  STRAIN=H37Rv;
RX  MEDLINE=9825987; PubMed=9634230; DOI=10.1038/311159;
RX  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA  Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA  Tekaiia F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA  Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA  Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA  Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA  Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA  Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT  "Deciphering the biology of Mycobacterium tuberculosis from the
RT  complete genome sequence.";
RL  Nature 393:537-544(1998).
DR  EMBL; BX842584; CAE55649.1; -; Genomic_DNA.
DR  Tuberculist; Rv3892c; -.
DR  InterPro; IPR000030; Microbac_PPE.
DR  Pfam; PF00823; PPE; 1.
KW  Complete proteome.
SQ  SEQUENCE 399 AA; 38852 MW; 30C2EEFB07455A91 CRC64;

Query Match          6.8%; Score 8; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  18 GGVLAALA 25
    |||||
Db  160 GGVLAALA 167

RESULT 315
Q87JT0_VIBPA

```

```
ID O87JTO_VIBPA PRELIMINARY; PRT; 401 AA.
AC O87JTO;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Multidrug resistance protein D.
DS OrderedLocusNames=VPA0168;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
OX NCBI_TaxID=670;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RMID 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yaenaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; BA000032; BAC61511.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0015904; P:tetracycline transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004812; Efflux_Bcr_Cf1A.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1.
DR PRINTS; PR01036; TCRTEB.
DR TIGRFAMs; TIGR00710; efflux_Bcr_Cf1A; 1.
DR PROSITE; PS00850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 401 AA; 42413 MW; 1D49618D381248B8 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVLA 22
Db 234 VLLGGVLA 241
|||||
|||||

RESULT 316
CBPAL_PIG STANDARD; PRT; 419 AA.
AC P09954; Q9TV85;
DT 01-MAR-1989 (Rel. 10, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Carboxypeptidase A1 precursor (EC 3.4.17.1).
GN Name=CPA1; Synonyms=CPA;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Pancreas;
RX MEDLINE=99192816; PubMed=10092856;
RA Darnis S., Juge N., Marino C., Aviles F.X., Puigserver A.,
RA Chaix J.-C., Guo X.-J.;
RT "Cloning, sequencing and functional expression of a cDNA encoding
RT porcine pancreatic preprocarboxypeptidase A1.";
RL Eur. J. Biochem. 259:719-725(1999).
RN [2]
RP PROTEIN SEQUENCE OF 17-110.
RX MEDLINE=87100171; PubMed=3801014;

RA Vendrell J., Aviles F.X., Genesca E., San Segundo B., Soriano F.,
RA Mendez E.;
RT "Primary structure of the activation segment of procarboxypeptidase A
RT from porcine pancreas.";
RL Biochem. Biophys. Res. Commun. 141:517-523(1986).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2 ANGSTROMS) OF 17-419.
RX MEDLINE=92194312; PubMed=1548696;
RA Guasch A., Coll M., Aviles F.X., Huber R.;
RT "Three-dimensional structure of porcine pancreatic procarboxypeptidase
RT A. A comparison of the A and B zymogens and their determinants for
RT inhibition and activation.";
RL J. Mol. Biol. 224:141-157(1992).
CC -!- CATALYTIC ACTIVITY: Release of a C-terminal amino acid, but little
CC or no action with -Asp, -Glu, -Arg, -Lys or -Pro.
CC -!- COFACTOR: Binds 1 zinc ion per subunit.
CC -!- SUBUNIT: Monomer.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the peptidase M14 family.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
DR EMBL; AF076222; AAD17690.1; -; mRNA.
DR PIR; A25833; A25833
DR PDB; 1PCA; X-ray; @=-.
DR SMR; P09954; 17-417.
DR MEROPS; M14.001; -.
DR InterPro; IPR000834; Peptidase M14.
DR InterPro; IPR003146; Prot inh M14A.
DR Pfam; PF00246; Peptidase M14; 1.
DR Pfam; PF02244; Propep M14; 1.
DR PRINTS; PR00765; CRBOXYPTASEA.
DR PROSITE; PS00132; CARBOXYPEPT_ZN_1; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; 1.
KW 3D-structure; Carboxypeptidase; Direct protein sequencing; Hydrolase;
Metal-binding; Metalloprotease; Protease; Signal; Zinc; Zymogen.
FT SIGNAL 1 16
FT PROPEP 17 110 Activation peptide.
FT CHAIN 111 419 Carboxypeptidase A1.
FT ACT_SITE 358 358 Proton donor.
FT ACT_SITE 380 380 Nucleophile.
FT METAL 179 179 Zinc.
FT METAL 182 182 Zinc.
FT METAL 306 306 Zinc.
FT METAL 248 271 Zinc.
FT TURN 21 22
FT STRAND 24 28
FT HELIX 33 42
FT TURN 43 44
FT HELIX 46 48
FT STRAND 51 54
FT TURN 59 60
FT STRAND 63 67
FT HELIX 69 71
FT HELIX 72 81
FT TURN 82 83
FT STRAND 86 90
FT HELIX 93 106
FT TURN 107 107
SQ SEQUENCE 419 AA; 47235 MW; 84B4CB557B714FC1 CRC64;

Query Match 6.8%; Score 8; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVLA 22
Db 9 VLLGGVLA 16
|||||
|||||
```

```

RESULT 317
Q9H833 HALSA
ID Q9H833 HALSA PRELIMINARY; PRT; 420 AA.
AC Q9H833;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Multidrug resistance protein homolog.
GN Name=yf02; OrderedLocustNames=VNG0427G;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
ON NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahaitas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Bailiga N.S., Thorson V., Shrogha J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jabloncki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Ienberger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE004998; AAG18975.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS00850; MFS_1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 420 AA; 42604 MW; DB3A1C24F532A0CD CRC64;

Query Match 6.8%; Score 8; DB 2; Length 420;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALA 25
Db 389 GGVLAALA 396
|||||

RESULT 318
Q68344 9HEPC
ID Q68344 9HEPC PRELIMINARY; PRT; 425 AA.
AC Q68344;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HCV-KI;
RA Paik S.H., Yang J.M.;
RT "Hepatitis C virus genome, complete NS4 and part of NS3 and NS5
sequence.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U26687; AAA79971.1; -; Genomic_RNA.
DR HSP; P27958; IHEI.
DR SMR; Q68344; 1-51.

```

```

DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
KW Polypeptide.
FT CHAIN <1 9 NS3 protein.
FT CHAIN 10 406 NS4 protein.
FT CHAIN 407 >425 NS5 protein.
FT NON_TER 1 1
FT NON_TER 425 425
SQ SEQUENCE 425 AA; 45366 MW; 1A2474932E0EB262 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 425;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 42 CMSADLEV 49
|||||

RESULT 319
Q9QQP2 9HEPC
ID Q9QQP2 9HEPC PRELIMINARY; PRT; 449 AA.
AC Q9QQP2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE NS3 helicase (Fragment).
GN Name=pol;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
ON NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Xu J.J., Wang H.T., Wang T., Qi L.Q.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF156794; AAD45674.1; -; Genomic_RNA.
DR PIR; A61196; A61196.
DR PIR; PS0329; PS0329.
DR HSP; P26664; IHEI.
DR SMR; Q9QQP2; 1-449.
DR GO; GO:0005224; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; Helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR KW Helicase.
FT NON_TER 1 1
FT NON_TER 449 449
SQ SEQUENCE 449 AA; 48016 MW; 316A6BD7F55C87E2 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 440 CMSADLEV 447
|||||

RESULT 320
Q9WNM3 9HEPC
ID Q9WNM3 9HEPC PRELIMINARY; PRT; 449 AA.
AC Q9WNM3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```

Query Match	6.8%; Score 8; DB 2; Length 449;
Best Local Similarity	100.0%; Pred. No. 57;
Matches	8; Conservative 0; Mismatches 0; Indels
QY	2 CMSADLEV 9
DB	440 CMSADLEV 447
RESULT 323	
Q9QOP3_9HEPC	
ID	Q9QOP3_9HEPC PRELIMINARY; PRT; 449 AA.
AC	Q9QOP3;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	N53 helicase (Fragment).
GN	Name=pol;
OS	Hepatitis C virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Flavi
OC	Hepacivirus.
OX	NCBI_TaxID=11103;
OX	[1]
RP	NUCLEOTIDE SEQUENCE.
RA	Xu J.J., Wang H.T., Wang T., Qi L.Q.;
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF156793; AAD45673.1; -; Genomic_RNA.
DR	HSSP; P26664; 1HEI.
DR	SMR; Q9QOP3; 1-449.
DR	GO; GO:0005524; P.ATP binding; IEA.
DR	GO; GO:0008026; F.ATP-dependent helicase activity; IEA.
DR	GO; GO:0003676; F.nucleic acid binding; IEA.
DR	InterPro; IPR001410; DEAD.
DR	InterPro; IPR011545; DEAD/DEAH_N.
DR	InterPro; IPR001650; Helicase_C.
DR	Pfam; PF00271; Helicase_C; 1.
DR	SMART; SM00487; DEXDC; 1.
KW	Helicase.
FT	NON TER 1
FT	NON TER 449 449
FT	SEQUENCE 449 AA; 47880 MW; ED22661D313968AA CRC64;
QY	2 CMSADLEV 9
DB	440 CMSADLEV 447
RESULT 323	
Q4NTJ3_9DEL	
ID	Q4NTJ3_9DEL PRELIMINARY; PRT; 559 AA.
AC	Q4NTJ3;
DT	13-SEP-2005 (TrEMBLrel. 31, Created)
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE	Anthraxlate phosphoribosyl transferase (EC 2.4.2.18).
GN	ORFNames=AdenDRAFT_2049;
OS	Anaeromyxobacter dehalogenans 2CP-C.
OC	Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcal
OC	Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX	NCBI_TaxID=290397;
OX	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=2CP-C;
RG	US DOE Joint Genome Institute (JGI-PGF);
RA	Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., G.
RA	Hannon N., Israeli S., Pittluck S., Richardson P.;
RA	"Sequencing of the draft genome assembly of Anaeromyxobac
RT	dehalogenans 2CP-C.";
RT	Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RL	Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.

```

RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN-2CP-C.
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000017; EAL78934.1; -; Genomic_DNA.
KW Glycoyltransferase; Transferase.
SQ SEQUENCE 559 AA; 57723 MW; F048B229F85726EC CRC64;

Query Match 6.8%; Score 8; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLAAL 24
DB 425 LGGVLAAL 432

RESULT 324
Q4H6D4_9DETO
ID Q4H6D4_9DETO PRELIMINARY; PRT; 560 AA.
AC Q4H6D4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Similar to dehydrogenase precursor.
GN OAFNames=DgeODRAFT_2962;
OS Deinococcus geothermalms DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Iserani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalms
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000015; EAL81818.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL
FT SIGNAL
SQ SEQUENCE 560 AA; 59616 MW; 216C452B2597B98A CRC64;

Query Match 6.8%; Score 8; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAAY 27
DB 85 VLAALAAY 92

RESULT 325
Q8EH53_SHEON
ID Q8EH53_SHEON PRELIMINARY; PRT; 592 AA.
AC Q8EH53;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein SO1377.
GN OrderedLocusNames=SO1377;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MR-1;
RX MSDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nealsen K.H., Fraser C.M.;
RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015581; AAN54442.1; -; Genomic_DNA.
DR TIGR; SO1377; -.
GO GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR001107; Band_7.
DR Pfam; PF01145; Band_7; 1.
DR SMART; SM00244; PHB; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 592 AA; 65642 MW; 762DDEDDC65647D2 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 71;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 QQQAVIEP 102
DB 581 QQQAVIEP 588

RESULT 326
Q4JME8_9HEPC
ID Q4JME8_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JME8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Names=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Saulea S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068207; AAY84772.1; -; Genomic_RNA.
FT NON TER
FT NON TER
SQ SEQUENCE 631 AA; 67064 MW; B6A49414FE0CE3B0 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;

```


Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 327
Q4JME9_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JME9_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JME9;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RL transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068206; AAY84771.1; -; Genomic_RNA.
FT NON_TER 1 631
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67036 MW; ECC07ED813122C84 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 328
Q4JMF0_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMF0_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF0;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RL transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068205; AAY84770.1; -; Genomic_RNA.
FT NON_TER 1 631
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67094 MW; CFFACEDCF5037EE CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9

Db 622 CMSADLEV 629
|||||

RESULT 329
Q4JMF1_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMF1_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF1;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RL transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068204; AAY84769.1; -; Genomic_RNA.
FT NON_TER 1 631
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67050 MW; BCA49414E40CE877 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 330
Q4JMF2_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMF2_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF2;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RL transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068203; AAY84768.1; -; Genomic_RNA.
FT NON_TER 1 631
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67080 MW; B6A49414F88646AA CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

```
RESULT 331
Q4JMF3_9HEPC
ID Q4JMF3_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68201; AY84764.1; -; Genomic_RNA.
FT NON_TER 1
FT NON_TER 631
SQ SEQUENCE 631 AA; 67064 MW; B6A49414FE0CE3B0 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 332
Q4JMF4_9HEPC
ID Q4JMF4_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68201; AY84764.1; -; Genomic_RNA.
FT NON_TER 1
FT NON_TER 631
SQ SEQUENCE 631 AA; 67150 MW; AABBF92420612D87 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 333
Q4JMF5_9HEPC
ID Q4JMF5_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68201; AY84764.1; -; Genomic_RNA.
FT NON_TER 1
FT NON_TER 631
SQ SEQUENCE 631 AA; 67062 MW; B6A0867B8318D75D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
```

```
RESULT 334
Q4JMF6_9HEPC
ID Q4JMF6_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68199; AY84764.1; -; Genomic_RNA.
FT NON_TER 1
FT NON_TER 631
SQ SEQUENCE 631 AA; 67062 MW; B6A0867B8318D75D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 335
Q4JMF7_9HEPC
ID Q4JMF7_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
```

```
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068196; AAY84761.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67163 MW; 2E797E1D93B906F1 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 336
Q4JMF8_9HEPC
ID Q4JMF8_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF8;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068197; AAY84762.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67092 MW; B6A49414FE0A3587 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 337
Q4JMF9_9HEPC
ID Q4JMF9_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMF9;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068198; AAY84763.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67062 MW; B6A0867B8318D75D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 338
Q4JMG0_9HEPC
ID Q4JMG0_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG0;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068195; AAY84760.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67064 MW; B6A49414FE0CE3B0 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 339
Q4JMG1_9HEPC
ID Q4JMG1_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG1;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
```

```
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068196; AAY84761.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67062 MW; B6A0867B8318D75D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 338
Q4JMG0_9HEPC
ID Q4JMG0_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG0;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068195; AAY84760.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67064 MW; B6A49414FE0CE3B0 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629

RESULT 339
Q4JMG1_9HEPC
ID Q4JMG1_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG1;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
```

```
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68192; AAY84757.1; -; Genomic_RNA.
FT NON_TER 1 1
FT NON_TER 631 631
SQ SEQUENCE 631 AA; 67094 MW; 26773E1451F91EB4 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db |||||||
622 CMSADLEV 629

RESULT 340
Q4JMG2_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMG2_9HEPC
AC Q4JMG2;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68193; AAY84758.1; -; Genomic_RNA.
FT NON_TER 1 1
FT NON_TER 631 631
SQ SEQUENCE 631 AA; 67080 MW; 53899E978B79A68D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db |||||||
622 CMSADLEV 629

RESULT 341
Q4JMG3_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMG3_9HEPC
AC Q4JMG3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
```

```
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68192; AAY84757.1; -; Genomic_RNA.
FT NON_TER 1 1
FT NON_TER 631 631
SQ SEQUENCE 631 AA; 67033 MW; 5DC536B8424E185B CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db |||||||
622 CMSADLEV 629

RESULT 342
Q4JMG4_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMG4_9HEPC
AC Q4JMG4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68191; AAY84756.1; -; Genomic_RNA.
FT NON_TER 1 1
FT NON_TER 631 631
SQ SEQUENCE 631 AA; 67062 MW; B6A0867B8318D75D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CMSADLEV 9
Db |||||||
622 CMSADLEV 629

RESULT 343
Q4JMG5_9HEPC PRELIMINARY; PRT; 631 AA.
ID Q4JMG5_9HEPC
AC Q4JMG5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCB1_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
```

RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068189; AAY84755.1; -, Genomic_RNA.
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67076 MW; D2323F5014F91EE5 CRC64;
Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 344
Q4JMG6_9HEPC
ID Q4JMG6_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG6;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RT "Evolution of the gene encoding the complete nonstructural 3 protein
(NS3) of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068189; AAY84754.1; -, Genomic_RNA.
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67076 MW; D2323F5014F91EE5 CRC64;
Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 345
Q4JMG7_9HEPC
ID Q4JMG7_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG7;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RT "Evolution of the gene encoding the complete nonstructural 3 protein
(NS3) of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068189; AAY84753.1; -, Genomic_RNA.
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67050 MW; 853857B86F42A9BF CRC64;
Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 346
Q4JMG8_9HEPC
ID Q4JMG8_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG8;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RT "Evolution of the gene encoding the complete nonstructural 3 protein
(NS3) of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068187; AAY84752.1; -, Genomic_RNA.
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67122 MW; D2323F5014F8F361 CRC64;
Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 347
Q4JMG9_9HEPC
ID Q4JMG9_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG9;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RT "Evolution of the gene encoding the complete nonstructural 3 protein
(NS3) of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068186; AAY84751.1; -, Genomic_RNA.

RT transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068188; AAY84753.1; -, Genomic_RNA.
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67050 MW; 853857B86F42A9BF CRC64;
Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 346
Q4JMG8_9HEPC
ID Q4JMG8_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG8;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RT "Evolution of the gene encoding the complete nonstructural 3 protein
(NS3) of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068187; AAY84752.1; -, Genomic_RNA.
FT NON_TER 1 631
SQ SEQUENCE 631 AA; 67122 MW; D2323F5014F8F361 CRC64;
Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred.No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CMSADLEV 9
Db 622 CMSADLEV 629
RESULT 347
Q4JMG9_9HEPC
ID Q4JMG9_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMG9;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RT "Evolution of the gene encoding the complete nonstructural 3 protein
(NS3) of Hepatitis C virus (HCV) during chronic infection and after
transmission to a new host.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ068186; AAY84751.1; -, Genomic_RNA.

```
FT  NON TER      1      1
FT  NON_TER     631    631
SQ  SEQUENCE    631 AA; 67031 MW; 1310F1A0DC473847 CRC64;

Query Match      6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 CMSADLEV 9
Db  622 CMSADLEV 629

RESULT 348
Q4JMH0_9HEPC
ID  Q4JMH0_9HEPC PRELIMINARY; PRT; 631 AA.
AC  Q4JMH0;
DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE  Nonstructural protein 3 (Fragment).
GN  Name=NS3;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.
OX  NCBI_TaxID=11103;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA  Otero T., Esteban J.I., Esteban R., Guardia J.;
RT  "Evolution of the gene encoding the complete nonstructural 3 protein
RT  (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT  transmission to a new host.";
RL  Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR  EMBL; DQ068185; AAY84750.1; -; Genomic_RNA.
FT  NON TER      1      1
FT  NON_TER     631    631
SQ  SEQUENCE    631 AA; 67050 MW; B4432414FFCA5380 CRC64;

Query Match      6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 CMSADLEV 9
Db  622 CMSADLEV 629

RESULT 349
Q4JMH1_9HEPC
ID  Q4JMH1_9HEPC PRELIMINARY; PRT; 631 AA.
AC  Q4JMH1;
DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE  Nonstructural protein 3 (Fragment).
GN  Name=NS3;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.
OX  NCBI_TaxID=11103;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA  Otero T., Esteban J.I., Esteban R., Guardia J.;
RT  "Evolution of the gene encoding the complete nonstructural 3 protein
RT  (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT  transmission to a new host.";
RL  Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR  EMBL; DQ068184; AAY84749.1; -; Genomic_RNA.
FT  NON TER      1      1
FT  NON_TER     631    631
SQ  SEQUENCE    631 AA; 67062 MW; 7D323F5014F90889 CRC64;
```

```
Query Match      6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 CMSADLEV 9
Db  622 CMSADLEV 629

RESULT 350
Q4JMH2_9HEPC
ID  Q4JMH2_9HEPC PRELIMINARY; PRT; 631 AA.
AC  Q4JMH2;
DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE  Nonstructural protein 3 (Fragment).
GN  Name=NS3;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.
OX  NCBI_TaxID=11103;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA  Otero T., Esteban J.I., Esteban R., Guardia J.;
RT  "Evolution of the gene encoding the complete nonstructural 3 protein
RT  (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT  transmission to a new host.";
RL  Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR  EMBL; DQ068183; AAY84748.1; -; Genomic_RNA.
FT  NON TER      1      1
FT  NON_TER     631    631
SQ  SEQUENCE    631 AA; 67062 MW; 7D323F5014F90889 CRC64;

Query Match      6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  2 CMSADLEV 9
Db  622 CMSADLEV 629

RESULT 351
Q4JMH3_9HEPC
ID  Q4JMH3_9HEPC PRELIMINARY; PRT; 631 AA.
AC  Q4JMH3;
DT  13-SEP-2005 (TrEMBLrel. 31, Created)
DT  13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE  Nonstructural protein 3 (Fragment).
GN  Name=NS3;
OS  Hepatitis C virus.
OC  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC  Hepacivirus.
OX  NCBI_TaxID=11103;
RN  [1]
RP  NUCLEOTIDE SEQUENCE.
RA  Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA  Otero T., Esteban J.I., Esteban R., Guardia J.;
RT  "Evolution of the gene encoding the complete nonstructural 3 protein
RT  (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT  transmission to a new host.";
RL  Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR  EMBL; DQ068182; AAY84747.1; -; Genomic_RNA.
FT  NON TER      1      1
FT  NON_TER     631    631
SQ  SEQUENCE    631 AA; 67062 MW; B6A0867B8318D75D CRC64;

Query Match      6.8%; Score 8; DB 2; Length 631;
```

Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 352
Q4JMH4_9HEPC
ID Q4JMH4_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMH4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A-10/1998;
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68181; AAY84745.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 67076 MW; D2323F5014F91EE5 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 353
Q4JMH5_9HEPC
ID Q4JMH5_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMH5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A-10/1997;
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68180; AAY84745.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; B6A6C2A762C2725D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 354
Q4JMH6_9HEPC
ID Q4JMH6_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMH6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A-1/1996;
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68179; AAY84744.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 66981 MW; B20999B72CDC72ED CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

RESULT 355
Q4JMH7_9HEPC
ID Q4JMH7_9HEPC PRELIMINARY; PRT; 631 AA.
AC Q4JMH7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nonstructural protein 3 (Fragment).
GN Name=NS3;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A-10/1993;
RA Quer J., Cos J., Sauleda S., Ocana L., Murillo P., Martell M.,
RA Otero T., Esteban J.I., Esteban R., Guardia J.;
RT "Evolution of the gene encoding the complete nonstructural 3 protein
RT (NS3) of Hepatitis C virus (HCV) during chronic infection and after
RT transmission to a new host."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; DQ68178; AAY84743.1; -; Genomic_RNA.
FT NON TER 1
FT NON TER 1
FT NON TER 631
SQ SEQUENCE 631 AA; 1F633E71F6C3665E CRC64;

Query Match 6.8%; Score 8; DB 2; Length 631;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
|||||
Db 622 CMSADLEV 629

```

QY      2 CMSGADLEV 9
DB      622 CMSGADLEV 629

RESULT 356
Q4ZRO2_PRESY PRELIMINARY; PRT; 657 AA.
AC      Q4ZRO2;
DT      13-SEP-2005 (TRMBLrel. 31, Created)
DT      13-SEP-2005 (TRMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TRMBLrel. 31, Last annotation update)
DE      Cytochrome c-type biogenesis protein Ccmf precursor.
GN      ORFNames=Psy3388;
OS      Pseudomonas syringae pv. syringae B728a.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=205918;
[1]
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=B728a;
RG      DOE Joint Genome Institute;
RA      Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
RA      Nolan M., Goltzman E., Thiel J., Malfatti S., Lapidis A., Detter J.C.,
RA      Land M., Richardson P.M., Kyripides N.C., Ivanova N.;
RT      "Comparison of two complete genome sequences of Pseudomonas syringae
RL      pv. syringae B728a and pv. tomato DC3000."
RL      Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
[2]
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=B728a;
RA      Loper J.;
RL      Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[3]
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=B728a;
RA      Fell H., Fell W.S., Lindow S.E.;
RL      Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR      EMBL; CP000075; AY38420.1; -; Genomic_DNA.
DR      InterPro; IPR003567; Cyt_c_biog.
DR      InterPro; IPR002541; CytC asm.
DR      InterPro; IPR003568; CytC_biog_Ccmf.
DR      Pfam; PF01578; Cytochrom_C_asm; 1.
DR      PRINTS; PR01410; CCBIOGENESIS.
DR      PRINTS; PR01411; CCMFBIOGNIS.
DR      TIGRFAMs; TIGR00353; nrf; 1.
KW      Signal.
FT      SIGNAL.
SQ      SEQUENCE 657 AA; 71736 MW; FBDP7F15AB32F101 CRC64;
        6.8%; Score 8; DB 2; Length 657;
Query Match
Best Local Similarity 100.0%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 LGGVLAAL 24
DB      627 LGGVLAAL 634

RESULT 357
Q87Z08_PRESM PRELIMINARY; PRT; 657 AA.
AC      Q87Z08;
DT      01-JUN-2003 (TRMBLrel. 24, Created)
DT      01-JUN-2003 (TRMBLrel. 24, Last sequence update)
DT      01-MAR-2004 (TRMBLrel. 26, Last annotation update)
DE      Cytochrome c-type biogenesis protein Ccmf.
GN      Name=ccmf; OrderedLocNames=PSPT03630;
OS      Pseudomonas syringae (pv. tomato).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC      Pseudomonadaceae; Pseudomonas.
OX      NCBI_TaxID=323;

```

```

[1]
RN      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC      STRAIN=DC3000;
RX      MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA      Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA      Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA      Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA      Nelson W.C., Davidsten T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA      Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA      Uytterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA      Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA      Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA      Bender C.L., White O., Fraser C.M., Collmer A.;
RT      "The complete genome sequence of the Arabidopsis and tomato pathogen
RT      Pseudomonas syringae pv. tomato DC3000."
RL      Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR      EMBL; AB016853; AA057101.1; -; Genomic_DNA.
DR      TIGR; PSPT03630; -;
DR      GO; GO:0016020; C:membrane; IEA.
DR      GO; GO:0015232; F:heme transporter activity; IEA.
DR      GO; GO:0008535; P:cytochrome c oxidase complex assembly; IEA.
DR      GO; GO:0017004; P:cytochrome complex assembly; IEA.
DR      GO; GO:0015886; P:heme transport; IEA.
DR      InterPro; IPR002541; CytC asm.
DR      InterPro; IPR003568; CytC_biog_Ccmf.
DR      InterPro; IPR003570; CytC_biog_Nrfe.
DR      InterPro; IPR003567; Cyt_c_biog.
DR      Pfam; PF01578; Cytochrom_C_asm; 1.
DR      PRINTS; PR01410; CCBIOGENESIS.
DR      PRINTS; PR01411; CCMFBIOGNIS.
DR      PRINTS; PR01413; NRPEBIOGNIS.
DR      TIGRFAMs; TIGR00353; nrf; 1.
KW      Complete proteome.
SQ      SEQUENCE 657 AA; 71905 MW; 1735828D870C2DC8 CRC64;
        6.8%; Score 8; DB 2; Length 657;
Query Match
Best Local Similarity 100.0%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      17 LGGVLAAL 24
DB      627 LGGVLAAL 634

RESULT 358
Q4PH79_USTWA PRELIMINARY; PRT; 707 AA.
AC      Q4PH79;
DT      13-SEP-2005 (TRMBLrel. 31, Created)
DT      13-SEP-2005 (TRMBLrel. 31, Last sequence update)
DT      13-SEP-2005 (TRMBLrel. 31, Last annotation update)
DE      Hypothetical protein.
GN      ORFNames=UM00534.1;
OS      Ustilago maydis 521.
OC      Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes;
OC      Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OX      NCBI_TaxID=237631;
[1]
RN      NUCLEOTIDE SEQUENCE.
RC      STRAIN=521;
RA      Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA      Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA      Arachchi H., Armbruster J., Bachanteang P., Baldwin J., Barry A.,
RA      Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA      Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA      Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA      Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA      David R., Dawoe T., Degray S., Dodge S., Doolley K., Dorje P.,
RA      Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA      Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA      Fitzgerald M., Foley K., Gage D., Galaqan J., Gearin G., Gnerre S.,
RA      Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA      Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,

```


RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby B., Iliev I.,
RA Jaffe D., Jones C., Kamat A., Kanat A., Kamysseles M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lana D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokysang T., Lokysang Y., Lucien O.,
RA Lui A., Ma L., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabelli R., Maru K., Matthews C., Mauceli E.,
RA McCarthy J., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Minova T., Mikkelson T., Mlenga V., Moru K.,
RA Mores J., Mullain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamla T., Teomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of *Ustilago maydis*.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD0100012; EAK80986.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 707 AA; 74663 MW; 76B3C86CC7D661D CRC64;

Query Match 6.8%; Score 8; DB 2; Length 707;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAALA 25
|||||
Db 306 GGVLAALA 313

RESULT 359
Q5YPX2 NOCPA PRELIMINARY; PRT; 713 AA.
AC Q5YPX2
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=nfa49170;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of *Nocardia farcinica* IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD59769.1; -; Genomic_DNA.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR000594; Thif_domain.
DR Pfam; PF00899; Thif; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 713 AA; 76561 MW; 21EE386F5E306793 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 713;
Best Local Similarity 100.0%; Pred. No. 83;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAALAA 26
|||||
Db 485 GVLAAALAA 492

RESULT 360
Q5AYN4 EMENI PRELIMINARY; PRT; 760 AA.
AC Q5AYN4
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Hypothetical protein.
GN ORENAMES=AN6596.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gherre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamat A., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of *Aspergillus nidulans*.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD01000110; EAA58125.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 760 AA; 85435 MW; 4F8AD163A2D75AF6 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 760;
Best Local Similarity 100.0%; Pred. No. 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 ALAAYCLIS 30
|||||
Db 180 ALAAYCLIS 187

RESULT 361
P90192_9HEPC PRELIMINARY; PRT; 3010 AA.
ID P90192_9HEPC
AC P90192;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus.
 RN NCBI_TaxID=11103;
 RP [1]
 RC NUCLEOTIDE SEQUENCE.
 RX STRAIN=HCV-1b;
 RA MEDLINE=95340824; PubMed=7542279;
 RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
 RA Yamamoto C., Izumi N., Marumo F., Sato C.;
 RT "Comparison of full-length sequences of interferon-sensitive and
 RT resistant hepatitis C virus 1b.";
 RL J. Clin. Invest. 96:224-230(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HCV-1b;
 RA Enomoto N.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RA "CDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93224886; PubMed=8385694;
 RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
 RA Dusheiko G., Saeed A.A., Holmes E.C.;
 RA "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity.";
 RL J. Gen. Virol. 74:661-668(1993).
 DR EMBL; D50483; BAA09074.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.
 DR PIR; PQ0804; PQ0804.
 DR PIR; P80329; P80329.
 DR HSP; Q8JYS1; ICWX.
 DR SMR; P90192; 1029-1657, 2008-2170, 2420-2949.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV_capsid.
 DR InterPro; IPR002521; HCV_core.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NS1.
 DR InterPro; IPR000745; HCV_NS4A.
 DR InterPro; IPR001490; HCV_NS4B.
 DR InterPro; IPR002868; HCV_NS5A.
 DR InterPro; IPR002166; HCV_RdRp.
 DR InterPro; IPR001650; Helicase_C.
 DR InterPro; IPR002518; Pept_U39_HCV_NS2.
 DR InterPro; IPR004109; Peptidase_S29.
 DR InterPro; IPR007095; RNA_pol_DS_Ps.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.

DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00271; Helicase_C; 1.
 DR Pfam; PF00998; RdRp_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW Envelope protein; Polyprotein.
 FT CHAIN 1 191 core protein.
 FT CHAIN 192 383 E1.
 FT CHAIN 384 809 E2.
 FT CHAIN 810 1026 NS2.
 FT CHAIN 1027 1657 NS3.
 FT CHAIN 1658 1711 NS4A.
 FT CHAIN 1712 1972 NS4B.
 FT CHAIN 1973 2419 NS5A.
 FT CHAIN 2420 3010 NS5B.
 SQ SEQUENCE 3010 AA; 326761 MW; 6A8E26E6EC78B38D CRC64;
 Query Match 6.8%; Score 8; DB 2; Length 3010;
 Best Local Similarity 100.0%; Pred.No. 2.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 2 CMSADLEV 9
 Db 1648 CMSADLEV 1655
 RESULT 362
 P90195_9HEPC PRELIMINARY; PRT; 3010 AA.
 AC P90195;
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HCV-1b;
 RX MEDLINE=95340824; PubMed=7542279;
 RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
 RA Yamamoto C., Izumi N., Marumo F., Sato C.;
 RT "Comparison of full-length sequences of interferon-sensitive and
 RT resistant hepatitis C virus 1b.";
 RL J. Clin. Invest. 96:224-230(1995).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HCV-1b;
 RA Enomoto N.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=92044457; PubMed=1658209;
 RA Oshima M., Tsuchiya M., Yagasaki M., Orita T., Hasegawa M.,
 RA Tomonoh K., Kojima T., Hirata Y., Yamamoto O., Shio Y. et al.;
 RA "CDNA clones of Japanese hepatitis C virus genomes derived from a
 RT single patient show sequence heterogeneity.";
 RL J. Gen. Virol. 72:2805-2809(1991).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93224886; PubMed=8385694;
 RA Simmonds P., McOmish F., Yap P.L., Chan S.-W.W., Lin C.K.,
 RA Dusheiko G., Saeed A.A., Holmes E.C.;
 RA "Sequence variability in the 5' non-coding region of hepatitis C
 RT virus: identification of a new virus type and restrictions on sequence
 RT diversity.";
 RL J. Gen. Virol. 74:661-668(1993).
 DR EMBL; D50480; BAA09071.1; -; Genomic_RNA.
 DR PIR; A61196; A61196.
 DR PIR; PQ0246; PQ0246.

DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; Q8JYS1; 1CW.
DR SMT; P90195; 1029-1657, 2008-2170, 2420-2949.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR001506; HCV NS5a; 1.
DR InterPro; IPR00271; Helicase C; 1.
DR InterPro; IPR00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Envelope protein; Polyprotein.
FT CHAIN 1 191 core protein.
FT CHAIN 192 383 E1.
FT CHAIN 384 809 E2.
FT CHAIN 810 1026 NS2.
FT CHAIN 1027 1657 NS3.
FT CHAIN 1658 1711 NS4A.
FT CHAIN 1712 1972 NS4B.
FT CHAIN 1973 2419 NS5A.
FT CHAIN 2420 3010 NS5B.
SQ SEQUENCE 3010 AA; 326487 MW; D352C338857309E0 CRC64;

Query Match 6.8%; Score 8; DB 2; Length 3010;
Best Local Similarity 100.0%; Pred.No. 2.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CMSADLEV 9
Db 1648 CMSADLEV 1655

RESULT 363
O92530_9HEPC
ID O92530_9HEPC PRELIMINARY; PRT; 3013 AA.
AC O92530;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=VN235;
RX MEDLINE=98378034; PubMed=9714232;
RA Tokita H., Okamoto H., Iizuka H., Kishimoto J., Tauda F., Miyakawa Y.,
RA Mayumi M.;
RT "The entire nucleotide sequences of three hepatitis C virus isolates
in genetic groups 7-9 and comparison with those in the other eight
genetic groups";
RL J. Gen. Virol. 79:1847-1857(1998).
DR EMBL; D84263; BAA32665.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CW.
DR SMT; O92530; 1027-1655, 2423-2988.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR001506; HCV NS5a; 1.
DR InterPro; IPR00271; Helicase C; 1.
DR InterPro; IPR00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3013 AA; 328197 MW; C9BE9C0231E86EAF CRC64;
Query Match 6.8%; Score 8; DB 2; Length 3013;
Best Local Similarity 100.0%; Pred.No. 2.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAYCL 29
Db 1666 AALAAAYCL 1673

RESULT 364
Q51DU1_ENTHI
ID Q51DU1_ENTHI PRELIMINARY; PRT; 61 AA.
AC Q51DU1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.

```

GN ORFNames=13.t00022;
OS Entamoeba histolytica HM-1:IMSS.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoef U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moulé S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitz E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sichevitz-Ponten T., Weber C., Singh U., Mukherjee C.,
RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000067; EAL50979.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 61 AA; 6160 MW; 01C331177097B888 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 34 VVIVGHI 40
Db 41 VVIVGHI 47

RESULT 365
ID Q98JE1_RHILO PRELIMINARY; PRT; 71 AA.
AC Q98JE1;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Msl1985 protein.
GN OrderedLocusName=msl1985;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=391;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MAFF303099;
RX MEDLINE=11082930; PubMed=11214969;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Matanabe A., Ikedawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyosawa K., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; BA000012; BAB49225.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 71 AA; 7566 MW; 9BDD37BF15915PAD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 58 VLAALAA 64

```

```

RESULT 366
ID Q9AF06_9ACTO PRELIMINARY; PRT; 76 AA.
AC Q9AF06;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Frankia sp. Cp11.
OC Plasmid pFQ12.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Frankineae; Frankiaceae; Frankia.
OX NCBI_TaxID=1856;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21429878; PubMed=11547880; DOI=10.1139/cjm-47-7-608;
RA John T.R., Rice J.M., Johnson J.D.;
RT "Analysis of pFQ12, a 22.4-kb Frankia plasmid.";
RL Can. J. Microbiol. 47:608-617(2001).
DR EMBL; AY027524; AAK20148.1; -; Genomic_DNA.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 76 AA; 7866 MW; 6664F80EDFA937EF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 10 VLAALAA 16

RESULT 367
ID GCH1_EUGGR STANDARD; PRT; 80 AA.
AC P51597;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE GTP cyclohydrolase I (EC 3.5.4.16) (GTP-CH-I) (Fragment).
OS Euglena gracilis.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Z / ATCC 12894;
RX MEDLINE=95352066; PubMed=7542887;
RA Maier J., Witter K., Guetlich M., Ziegler I., Werner T., Nimmemann H.;
RT "Homology cloning of GTP-cyclohydrolase I from various unrelated
RT eukaryotes by reverse-transcription polymerase chain reaction using a
RT general set of degenerate primers.";
RL Biochem. Biophys. Res. Commun. 212:705-711(1995).
CC -1- CATALYTIC ACTIVITY: GTP + 2 H(2)O = formate + 2-amino-4-hydroxy-6-
CC (erythro-1,2,3-trihydroxypropyl)-dihydropteridine triphosphate.
CC -1- ENZYME REGULATION: GTP shows a positive allosteric effect, and
CC tetrahydrobiopterin inhibits the enzyme activity (By similarity).
CC -1- PATHWAY: Cofactor biosynthesis; tetrahydrofolate biosynthesis; 2-
CC amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine diphosphate
CC from GTP: step 1.
CC -1- SUBUNIT: Homopolymer (By similarity).
CC -1- SIMILARITY: Belongs to the GTP cyclohydrolase I family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; Z49757; CAA89827.1; -; mRNA.
DR PIR; S54909; S54909.
DR HSP; P30793; 1PB1.

```

DR InterPro; IPR001474; GTP cyclohydrol.
 DR Pfam; PF01227; GTP_cyclohydrol; 1.
 DR ProDom; PD003330; GTP_cyclohydrol; 1.
 DR PROSITE; PS00859; GTP_CYCLOHYDROL_1_1; PARTIAL.
 DR PROSITE; PS00860; GTP_CYCLOHYDROL_1_2; 1.
 KW Allosteric enzyme; Hydrolase; Tetrahydrobiopterin biosynthesis.
 FT DISULFID 4 75 By similarity.
 FT NON_TER 1
 FT NON_TER 80
 SQ SEQUENCE 80 AA; 9097 MW; 8C46A15D01A7B8C7 CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 83 KGKVLGL 89
 Db 22 KGKVLGL 28
 RESULT 368
 Q8VJS5 MYCTU
 ID Q8VJS5 MYCTU PRELIMINARY; PRT; 82 AA.
 AC Q8VJS5
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=MT2068;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Winn M.L., Haft D.H.,
 RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL J. Bacteriol. 184:5479-5490(2002).
 RL EMBL; AE000516; AAK46346.1; -; Genomic_DNA.
 DR TIGR; MT2068; -.
 KW Hypothetical protein.
 SQ SEQUENCE 82 AA; 8822 MW; A8C6025292BE8CB0 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 82;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 20 VLAALAA 26
 Db 68 VLAALAA 74
 RESULT 369
 Q9Q3S2_9HEPC
 ID Q9Q3S2_9HEPC PRELIMINARY; PRT; 82 AA.
 AC Q9Q3S2;
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Polypeptide (Fragment).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.

OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21373979; PubMed=11481629; DOI=10.1053/jhep.2001.26635;
 RA Lin H.J., Seeff L.B., Barbosa L., Hollinger F.B.;
 RT "Occurrence of identical hypervariable region 1 sequences of hepatitis
 RT C virus in transfusion recipients and their respective blood donors:
 RT divergence over time";
 RL Hepatology 34:424-429(2001).
 DR EMBL; AF206351; AAF19863.1; -; Genomic_RNA.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0019031; C: viral envelope; IEA.
 DR InterPro; IPR002519; HCV_env.
 DR InterPro; IPR002531; HCV_NS1.
 DR Pfam; PF01539; HCV env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 KW Envelope protein; Polyprotein; Transmembrane.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 8597 MW; 87FEB7552B785F3D CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 82;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 GVLAALA 25
 Db 26 GVLAALA 32
 RESULT 370
 Q7NH49 GLOVI
 ID Q7NH49 GLOVI PRELIMINARY; PRT; 87 AA.
 AC Q7NH49;
 DT 01-MAR-2004 (TRENBLrel. 26, Created)
 DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Gsl2688 protein.
 GN OrderedLocusNames=gsl2688;
 OS Gloeobacter violaceus.
 OC Bacteria; Cyanobacteria; Gloeobacteridae; Gloeobacterales; Gloeobacter.
 OX NCBI_TaxID=33072;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=PCC 7421;
 RC MEDLINE=22977040; PubMed=14621292;
 RA Nakamura Y., Kaneo T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
 RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
 RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
 RT cyanobacterium that lacks thylakoids";
 RL DNA Res. 10:137-145(2003).
 DR EMBL; BA000045; BAC90629.1; -; Genomic_DNA.
 DR InterPro; IPR007367; DUF433.
 DR Pfam; PF04255; DUF433; 1.
 KW Complete proteome.
 SQ SEQUENCE 87 AA; 9651 MW; 7E6B117D8D0DCA10 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 42 LGKPAI 48
 Db 18 LGKPAI 24
 RESULT 371
 O39915_9HEPC
 ID O39915_9HEPC PRELIMINARY; PRT; 90 AA.
 AC O39915;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)

```
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Non-structural protein 4a/b (Fragment).
GN Name:NS4a/b;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98032593; PubMed=9365889;
RA DOI=10.1002/(SICI)1096-9071(199711)53:3<237::AID-JMV10>3.3.CO;2-P;
RA Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
RA Simmonds P.;
RT "Sequence analysis of hepatitis C virus variants producing discrepant
RT results with two different genotyping assays.";
RL J. Med. Virol. 53:237-244(1997).
DR EMBL; AF007520; AAB62971.1; -; Genomic_RNA.
DR HSP; P27958; 1A1R.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 90
FT NON_TER 90 90
SQ SEQUENCE 90 AA; 9742 MW; 2332FDS9CB38A04C9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 DEMEECS 66
DB 36 DEMEECS 42

RESULT 372
Q5SK97_THET8
ID Q5SK97_THET8 PRELIMINARY; PRT; 91 AA.
AC Q5SK97;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein THA0751.
GN OrderedLocNames=THA0751;
OS Thermus thermophilus (strain HB8 / ATCC 27634 / DSM 579).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=300852;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HB8;
RA Masui R., Kurokawa K., Nakagawa N., Tokunaga F., Koyama Y.,
RA Shibata T., Oshima T., Yokoyama S., Yasunaga T., Kuramitsu S.;
RT "Complete genome sequence of Thermus thermophilus HB8.";
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP008226; BAD70574.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 91 AA; 9846 MW; B7CD10CD7C0D4952 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LIGGVLA 22
DB 52 LIGGVLA 58

RESULT 373
Q532L0_9HEPC
ID Q532L0_9HEPC PRELIMINARY; PRT; 92 AA.
AC Q532L0;
```

```
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DB E2 protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RA Jouvencel A.-C., Neau D., Faure M., Neau M., Martinaud C., Legrand E.,
RA Dupon M., Ragnau J.-M., Fleury H., Lafon M.-E.;
RT "Hepatitis C virus variability in HIV co-infected patients, a study in
RT plasma and liver.";
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY793024; AAX22386.1; -; Genomic_RNA.
FT NON_TER 1 92
FT NON_TER 92 92
SQ SEQUENCE 92 AA; 9593 MW; C05C60D8CFDD5A62 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
DB 27 GVLAALA 33

RESULT 374
Q9HVFO_PSEAE
ID Q9HVFO_PSEAE PRELIMINARY; PRT; 96 AA.
AC Q9HVFO;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=PA4642;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004878; AAG08029.1; -; Genomic_DNA.
DR PIR; H83065; H83065.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 96 AA; 10800 MW; A85A332E40AB2FF8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 96;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 QAAPIIE 73
DB 82 QAAPIIE 88

RESULT 375
Q6ACG9_LEIXX
ID Q6ACG9_LEIXX PRELIMINARY; PRT; 102 AA.
```

```
AC Q6ACG9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Pyridine transhydrogenase.
GN OrderedLocusNames=LX22570;
OS Leifsonia xyl (subsp. xyl).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcales; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
RA de Oliveira J.C.P., Wood D., de Oliveira M.C., Miyaki C.Y.,
RA Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
RA Almeida N.F. Jr., Carrer H., Coutinho L.L., El-Dorri H.A.,
RA Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
RA Goldman G.H., Kimura E.T., Ferro E.S., Kuramae E.E., Lemos E.G.M.,
RA Lemos M.V.F., Mauro S.M.Z., Machado M.A., Marino C.L., Menck C.F.,
RA Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
RA Teai S.M., Zanca A.S., Simpson A.J.G., Brumley S.M., Setubal J.C.;
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
RT xyl subsp. xyl.";
RL Mol. Plant Microbe Interact. 17:827-836(2004).
RW EMBL; AE016822; AAT89924.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 102 AA; 10558 MW; D58C211E61596686 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 63 VLAALAA 69
|||||
63 VLAALAA 69

RESULT 376
Q9XP4_9HEPC PRELIMINARY; PRT; 102 AA.
AC Q9XP4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NON-structural protein NS4-GROUP I HCV-specific antigen C14-1
DE (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94245087; PubMed=7514558; DOI=10.1016/0270-9139(94)90226-7;
RA Tanaka T., Tsukiyama-Kohara K., Yamaguchi K., Yagi S., Tanaka S.,
RA Hasegawa A., Ohta Y., Hattori N., Kohara M.;
RT "Significance of specific antibody assay for genotyping of hepatitis C
RL Hepatology 19:1347-1353(1994).
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a_1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 102 102
SQ SEQUENCE 102 AA; 11389 MW; 267CA4C1F2F7F44E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 102;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
```

```
Db 38 PDKEVLY 44
|||||
38 PDKEVLY 44

RESULT 377
Q61GW9_CABR PRELIMINARY; PRT; 103 AA.
AC Q61GW9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG11016.
GN Name=CBG11016;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000052; CA65872.1; -; Genomic_DNA.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF03489; SapB_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 103 AA; 11244 MW; 5453B2B418EF45B5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 6 VLAALAA 12
|||||
6 VLAALAA 12

RESULT 378
Q6N7U3_RHOPA PRELIMINARY; PRT; 104 AA.
AC Q6N7U3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Possible transcriptional regulator, arsr family.
GN OrderedLocusNames=RPA2161;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison P.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RL photoautotrophic bacterium Rhodopseudomonas palustris.";
EL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572599; CA27602.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001845; HTH_Arsr.
DR Pfam; PF01022; HTH_5; 1.
DR PRINTS; PR00778; HTHARSR.
DR SMART; SM00418; HTH_Arsr; 1.
```


Qy 87 LGLLQRA 93
|||||
Db 53 LGLLQRA 59

RESULT 382
Q7NKD5 GLOVI
ID Q7NKD5 GLOVI PRELIMINARY; PRT; 122 AA.
AC Q7NKD5;
DT 01-MAR-2004 (TREMELrel. 26, Created)
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Grl1543 protein.
GN OrderedLocusNames=grl1543;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; BA000045; BAC9484.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 122 AA; 13628 MW; 70ABFE16C02B18D8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 122;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
|||||
Db 109 VLAALAA 115

RESULT 383
Q7NEL9 GLOVI
ID Q7NEL9 GLOVI PRELIMINARY; PRT; 124 AA.
AC Q7NEL9;
DT 01-MAR-2004 (TREMELrel. 26, Created)
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Gll3860 protein.
GN OrderedLocusNames=gll3860;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; BA000045; BAC91801.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 124 AA; 13902 MW; 88605D3313B506CF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 25
|||||
Db 22 GVLAAAL 28

RESULT 384
Q81592 9HEPC
ID Q81592 9HEPC PRELIMINARY; PRT; 125 AA.
AC Q81592;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE NS4 protein (fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]

NUCLEOTIDE SEQUENCE.
RX MEDLINE=93197128; PubMed=838385;
RA Sarashina T., Sakurai T., Watanabe Y., Kashima K., Suzuki T.,
RA Chiba J., Kita Y., Horiuchi T., Saito I., Miyamura T.;
RT "Nucleotide sequence of the hepatitis C virus genome from a patient
RT negative for anti-HCV by the first generation antibody assay.";
RL Nucleic Acids Res. 21:1037-1037(1993).
DR EMBL; D11353; BAA01956.1; -; Genomic_RNA.
DR PIR; S35629; S35629.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 125
FT NON_TER 125 125

SQ SEQUENCE 125 AA; 13680 MW; C5C8ADE45C7C1872 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 125;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
|||||
Db 58 PDKEVLY 64

RESULT 385
NUOA MYCBO
ID NUOA MYCBO STANDARD; PRT; 128 AA.
AC P65564; P95181;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE NADH-quinone oxidoreductase chain A (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain A) (NDH-1, chain A).
GN Name=nuoA; OrderedLocusNames=Mb3169;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AP2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is
CC believed to be menaquinone. Couples the redox reaction to proton

translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient (by similarity).

-1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: Belongs to the complex I subunit 3 family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL: BX248345; CAD95261.1; -; Genomic_DNA.
 InterPro: IPR000440; Oxidored_q4.
 PANTHER: PTHR11058; Oxidored_q4; 1.
 Pfam: PF00507; Oxidored_q4; 1.
 Complete proteome: NAD; Oxidoreductase; Quinone; Transmembrane.
 TRANSMEM 5 25 Potential.
 TRANSMEM 72 92 Potential.
 TRANSMEM 100 120 Potential.
 SSEQUENCE 128 AA; 13975 MW; DDF4FD3F77A8149A CRC64;

Query Match 5.9%; Score 7; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 9 VLAALAA 15

RESULT 386
 NUOA MYCTU STANDARD; PRT; 128 AA.
 ID NUOA MYCTU
 AC P65563; P95181;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DE 10-MAY-2005 (Rel. 47, Last annotation update)
 DE NADH-quinone oxidoreductase chain A (EC 1.6.99.5) (NADH dehydrogenase I, chain A) (NDH-1, chain A).
 DE I, chain A) (NDH-1, chain A).
 GN Names=nuoA; OrderedLocNames=Rv3145, MT3233; ORFNames=MYCY03A2.13c;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=H37Rv;
 RC MEDLINE=9825987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 Harris D.E., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III,
 Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
 Connor R., Davies R.M., Devlin K., Feldwell T., Gentles S., Hamlin N.,
 Holroyd S., Hornsby T., Jegels K., Krogh A., McLean J., Moule S.,
 Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence";
 RT Nature 393:537-544 (1998).
 RL Nature 393:537-544 (1998).
 [2]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=CDC 1551 / Oshkosh;
 RC MEDLINE=2206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
 Hickey E.K., Kolonay J.F., Nelson W.C., Unsay M.A., Ermolaeva M.D.,
 Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
 Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 Fraser C.M.;

Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.;
 J. Bacteriol. 184:5479-5490 (2002).

-1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-sulfur (Fe-S) centers, to quinones in the respiratory chain. The immediate electron acceptor for the enzyme in this species is believed to be menaquinone. Couples the redox reaction to proton translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient (By similarity).

-1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.

-1- SUBCELLULAR LOCATION: Integral membrane protein.

-1- SIMILARITY: Belongs to the complex I subunit 3 family.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL: BX842582; CAB06271.1; -; Genomic DNA.
 EMBL: AE000516; AAK47572.1; -; Genomic DNA.
 PIR: B70647; B70647.
 TIGR: MT3233; -;
 DR TubercuList; Rv3145; -;
 DR InterPro: IPR000440; Oxidored_q4.
 DR PANTHER: PTHR11058; Oxidored_q4; 1.
 DR Pfam: PF00507; Oxidored_q4; 1.
 Complete proteome: NAD; Oxidoreductase; Quinone; Transmembrane.
 TRANSMEM 5 25 Potential.
 TRANSMEM 72 92 Potential.
 TRANSMEM 100 120 Potential.
 SSEQUENCE 128 AA; 13975 MW; DDF4FD3F77A8149A CRC64;

Query Match 5.9%; Score 7; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 9 VLAALAA 15

RESULT 387
 Q956X0_9PLAT
 ID Q956X0_9PLAT PRELIMINARY; PRT; 128 AA.
 AC Q956X0;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome oxidase subunit I (Fragment).
 GN Name=COI;
 GN Choricotyle cf. chrysophryll.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Platyhelminthes; Monogenea; Polyopisthocotylea;
 OC Ncidiodophoridae; Choricotyle.
 OX NCBI_TaxID=146928;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC MEDLINE=21203684; PubMed=11306118; DOI=10.1016/S0020-7519(01)00114-X;
 RA Jovelin R., Justine J.L.;
 RT "Phylogenetic relationships within the polyopisthocotylean monogeneans (Platyhelminthes) inferred from partial 28S rDNA sequences.";
 RL Int. J. Parasitol. 31:393-401 (2001).
 CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory chain that catalyzes the reduction of oxygen to water. Subunits 1-3 form the functional core of the enzyme complex. CO I is the catalytic subunit of the enzyme. Electrons originating in cytochrome c are transferred via the copper A center of subunit 2 and heme A of subunit 1 to the bimetallic center formed by heme A3 and copper B (By similarity).
 CC -1- CATALYTIC ACTIVITY: 4 ferrocycytochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.

```
CC -!- PATHWAY: Respiratory chain; terminal step.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (By similarity).
CC -!- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
CC EMBL: AY009165; AAG33656.1; -; Genomic_DNA.
CC GO: GO:0019886; C:inner membrane; IEA.
CC GO: GO:0016021; C:integral to membrane; IEA.
CC GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
CC GO: GO:0005739; C:mitochondrion; IEA.
CC GO: GO:0004129; F:cytochrome-c oxidase activity; IEA.
CC GO: GO:0016491; F:oxidoreductase activity; IEA.
CC GO: GO:0006118; P:electron transport; IEA.
CC GO: GO:0006810; P:transport; IEA.
CC InterPro: IPR000883; COX1.
CC PANTHER: PTHR10422; COX1.
CC Pfam: PF00115; COX1.
CC PRINTS: PR01165; CYCOXIDASRI.
CC PROSITE: PS0855; COX1.
CC Copper; Electron transport; Heme; Inner membrane; Membrane;
KW Mitochondrion; Oxidoreductase; Respiratory chain; Transmembrane;
FT Transp.
FT NON_TER
FT TER
FT SEQUENCE 128 128
FT SEQUENCE 128 AA; 14014 MW; CFC495F4650454F6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 VIVGHIE 41
Db 8 VIVGHIE 14

RESULT 388
Q4NRG8_9DEL1 PRELIMINARY; PRT; 128 AA.
ID Q4NRG8_9DEL1 PRELIMINARY; PRT; 128 AA.
AC Q4NRG8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS ORFNames=AdenHRAFT_1318;
GN Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Dehalococcaceae; Myxococcales;
OC Cystobacteriaceae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Irani S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL: AAHD0100027; EAL78203.1; -; Genomic_DNA.
KW Hypothetical protein.
FT SEQUENCE 128 AA; 13243 MW; 4177FFC09B0750F2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 VIVGHIE 41
Db 8 VIVGHIE 14

RESULT 389
Q89SL2_BRAJA PRELIMINARY; PRT; 128 AA.
ID Q89SL2_BRAJA PRELIMINARY; PRT; 128 AA.
AC Q89SL2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome c2.
GN Name=Cy2; OrderedLocusNames=bll2388;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL: BA000040; BAC47653.1; -; Genomic_DNA.
DR HSSP: P54820; 1QL3.
DR GO: GO:0005489; F:electron transporter activity; IEA.
DR GO: GO:0020037; F:heme binding; IEA.
DR GO: GO:0006118; P:electron transport; IEA.
DR InterPro: IPR012282; Cytochrome_c_r.
DR InterPro: IPR003088; Cyt_c1.
DR InterPro: IPR002327; Cyt_c1AB.
DR InterPro: IPR009056; Cyt_c_monohaem.
DR Pfam: PF00034; Cytochrom_c1.
DR PRINTS: PR00604; CYTCRMECIAB.
DR PRODOM: PD000375; Cyt_C1AB; 1.
DR PROSITE: PS51007; CYTC; 1.
KW Complete proteome.
SQ SEQUENCE 128 AA; 13735 MW; 6747F2369DAA8C0A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 128;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 45 KPAIVPD 51
Db 65 KPAIVPD 71

RESULT 390
Q63P16_BURPS PRELIMINARY; PRT; 131 AA.
ID Q63P16_BURPS PRELIMINARY; PRT; 131 AA.
AC Q63P16;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BPS0132;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
```

RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
 RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
 RA Bentley S.D., Sebastian M., Thomson N.R., Bason N., Beacham I.R.,
 RA Brooks K., Brown K.A., Brown N.P., Challis G.L., Cherevach I.,
 RA Chillingworth T., Cronin A., Crosslet B., Davis P., Deshazer D.,
 RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagsels K.,
 RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 RA Rabinovitch E., Rutherford K., Sanders M., Simmonds M.,
 RA Songelvitai S., Stevens K., Tumapa S., Vesaratchavest M.,
 RA Whitehead S., Yeats C., Barrall B.G., Oyston P.C.F., Parkhill J.,
 RT "Genomic plasticity of the causative agent of melioidosis,
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
 DR EMBL; BX571966; CAH37577.1; -, Genomic_DNA.
 SW Complete proteome; Hypothetical protein.
 QX SEQUENCE 131 AA; 14808 MW; 83EP090384783DFA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 131;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
 |||||
 Db 22 GVLAALA 28

RESULT 391
 Q95618_9BASI
 ID Q95618_9BASI PRELIMINARY; PRT; 132 AA.
 AC Q95618;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=Cytb;
 OS Rhodotorula lactosa.
 OG Mitochondrion.
 OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
 OC Erythrobasidium clade; Rhodotorula.
 OX NCBI_TaxID=57693;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IPM 48516/IFO 1423;
 RX MEDLINE=21304469; PubMed=11411687;
 RA Bleswa S.K., Yokoyama K., Nishimura K., Miyaji M.;
 RT "Molecular phylogenetics of the genus Rhodotorula and related
 RT basidiomycetous yeasts inferred from the mitochondrial cytochrome b
 RT gene";
 RL Int. J. Syst. Evol. Microbiol. 51:1191-1199 (2001).
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex iii or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential
 CC coupled to ATP synthesis (By similarity).
 CC -1- COPACTOR: Binds 2 heme groups noncovalently (By similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome ci and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AB040633; BAB60985.1; -, Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom_B_C; 1.
 DR PROSITE; PS51003; Cytochrom_B_N; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 Qy 20 VLAALAA 26
 |||||
 Db 41 VLAALAA 47

KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 132 132
 SQ SEQUENCE 132 AA; 14737 MW; EE9B8F18D94E209F CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 20 VLAALAA 26
 |||||
 Db 41 VLAALAA 47
 RESULT 392
 Q956H5_9BASI
 ID Q956H5_9BASI PRELIMINARY; PRT; 132 AA.
 AC Q956H5;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=Cytb;
 OS Rhodospiridium paludigenum.
 OG Mitochondrion.
 OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
 OC Microbotryomycetidae; Sporidiobolales; Rhodospiridium.
 OX NCBI_TaxID=86838;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IPM 48537/IFO 10547;
 RX MEDLINE=21304469; PubMed=11411687;
 RA Bleswa S.K., Yokoyama K., Nishimura K., Miyaji M.;
 RT "Molecular phylogenetics of the genus Rhodotorula and related
 RT basidiomycetous yeasts inferred from the mitochondrial cytochrome b
 RT gene";
 RL Int. J. Syst. Evol. Microbiol. 51:1191-1199 (2001).
 CC -1- FUNCTION: Component of the ubiquinol-cytochrome c reductase
 CC complex (complex iii or cytochrome b-c1 complex), which is a
 CC respiratory chain that generates an electrochemical potential
 CC coupled to ATP synthesis (By similarity).
 CC -1- COPACTOR: Binds 2 heme groups noncovalently (By similarity).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome ci and the Rieske protein (By similarity).
 CC -1- SIMILARITY: Belongs to the cytochrome b family.
 DR EMBL; AB040646; BAB60998.1; -, Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom_B_C; 1.
 DR PROSITE; PS51003; Cytochrom_B_N; 1.
 DR PROSITE; PS51002; CYTB_NTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 Qy 20 VLAALAA 26
 |||||
 Db 41 VLAALAA 47

Query Match 5.9%; Score 7; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 2e+02; Indels 0; Gaps 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 20 VLAALAA 26
 |||||
 Db 41 VLAALAA 47

```

RESULT 393
Q95616_9BASI PRELIMINARY; PRT; 132 AA.
AC Q95616;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Rhodotorula marina.
OG Mitochondrion.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
OC Erythrobasidium clade; Rhodotorula.
OX NCBI_TaxID=106013;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 5761/IFO 0928;
RX MEDLINE=21304469; PubMed=11411687;
RA Biswa S.K., Yokoyama K., Nishimura K., Miyaji M.;
RT "Molecular phylogenetics of the genus Rhodotorula and related
RT basidiomycetous yeasts inferred from the mitochondrial cytochrome b
RT gene.";
RL Int. J. Syst. Evol. Microbiol. 51:1191-1199(2001).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC EMBL; AB040635; BAB60987.1; -; Genomic DNA.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0016020; C:membrane; IEA.
DR GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO:0005739; C:mitochondrion; IEA.
DR GO:0046872; F:metal ion binding; IEA.
DR GO:0016491; F:oxidoreductase activity; IEA.
DR GO:0006118; P:electron transport; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 14694 MW; 09169BD0268B119 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 41 VLAALAA 47

RESULT 394
Q95616_9BASI PRELIMINARY; PRT; 132 AA.
AC Q95616;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Erythrobasidium hasegawianum.
OG Mitochondrion.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
OC Erythrobasidium clade; Rhodotorula.
OX NCBI_TaxID=106001;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 48594/CBS 8076;
RX MEDLINE=21304469; PubMed=11411687;
RA Biswa S.K., Yokoyama K., Nishimura K., Miyaji M.;
RT "Molecular phylogenetics of the genus Rhodotorula and related
RT basidiomycetous yeasts inferred from the mitochondrial cytochrome b
RT gene.";
RL Int. J. Syst. Evol. Microbiol. 51:1191-1199(2001).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC EMBL; AB040635; BAB60987.1; -; Genomic DNA.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0016020; C:membrane; IEA.
DR GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO:0005739; C:mitochondrion; IEA.
DR GO:0046872; F:metal ion binding; IEA.
DR GO:0016491; F:oxidoreductase activity; IEA.
DR GO:0006118; P:electron transport; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 14694 MW; 09169BD0268B119 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 41 VLAALAA 47

```

```

OG Mitochondrion.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
OC Erythrobasidium clade; Erythrobasidium.
OX NCBI_TaxID=5414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 48186/IFO 1058;
RX MEDLINE=21304469; PubMed=11411687;
RA Biswa S.K., Yokoyama K., Nishimura K., Miyaji M.;
RT "Molecular phylogenetics of the genus Rhodotorula and related
RT basidiomycetous yeasts inferred from the mitochondrial cytochrome b
RT gene.";
RL Int. J. Syst. Evol. Microbiol. 51:1191-1199(2001).
CC -!- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (By similarity).
CC -!- COFACTOR: Binds 2 heme groups noncovalently (By similarity).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
CC EMBL; AB040635; BAB61017.1; -; Genomic DNA.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0016020; C:membrane; IEA.
DR GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO:0005739; C:mitochondrion; IEA.
DR GO:0046872; F:metal ion binding; IEA.
DR GO:0016491; F:oxidoreductase activity; IEA.
DR GO:0006118; P:electron transport; IEA.
DR GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR InterPro; IPR005797; Cytb_b6_N.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR Pfam; PF00033; Cytochrom B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 14733 MW; 9AE981ACBF0303ED CRC64;

Query Match 5.9%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 41 VLAALAA 47

RESULT 395
Q95616_9BASI PRELIMINARY; PRT; 132 AA.
AC Q95616;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Rhodotorula armeniaca.
OG Mitochondrion.
OC Eukaryota; Fungi; Basidiomycota; Urediniomycetes;
OC Erythrobasidium clade; Rhodotorula.
OX NCBI_TaxID=106001;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 48594/CBS 8076;
RX MEDLINE=21304469; PubMed=11411687;
RA Biswa S.K., Yokoyama K., Nishimura K., Miyaji M.;
RT "Molecular phylogenetics of the genus Rhodotorula and related
RT basidiomycetous yeasts inferred from the mitochondrial cytochrome b
RT gene.";

```

RL Int. J. Syst. Evol. Microbiol. 51:1191-1199(2001).
CC -I- FUNCTION: Component of the ubiquinol-cytochrome c reductase
CC complex (Complex III or cytochrome b-c1 complex), which is a
CC respiratory chain that generates an electrochemical potential
CC coupled to ATP synthesis (by similarity).
CC -I- COFACTOR: Binds 2 heme groups noncovalently (by similarity).
CC -I- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (by similarity).
CC -I- SIMILARITY: Belongs to the cytochrome b family.
CC EMBL; AB040615; BAB0967.1; -, Genomic DNA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrial; IEA.
DR GO; GO:0046872; F: metal ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; F: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR005797; Cytb_b6_C.
DR InterPro; IPR005798; Cytb_b6_N.
DR Pfam; PF00033; Cytochrom_B_N; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
DR PROSITE; PS51002; CYTB_NTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 132
FT NON_TER 132 132
SQ SEQUENCE 132 AA; 14733 MW; 95A1E9090AE2E4B1 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 132;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAALAA 26
Db 41 VLAAALAA 47

RESULT 396
Q68214_9HEPC
ID Q68214_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68214;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14250; AAC53939.1; -, Genomic RNA.
DR HSP; P27958; IHEI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 138
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15117 MW; 5995D942BCA5779 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAAALAA 26
Db 41 VLAAALAA 47

Qy 50 PDKEVLY 56
Db 70 PDKEVLY 76

RESULT 397
Q68219_9HEPC
ID Q68219_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68219;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14255; AAC53944.1; -, Genomic RNA.
DR HSP; P27958; IHEI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
FT NON_TER 1 138
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15100 MW; DC1B0A1DF0A837E5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 70 PDKEVLY 76

RESULT 398
Q68220_9HEPC
ID Q68220_9HEPC PRELIMINARY; PRT; 138 AA.
AC Q68220;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nonstructural protein (Fragment).
GN Name=ns4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2a;
RX MEDLINE=95146953; PubMed=7844535;
RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;
RT "Prevalence of hepatitis C virus sequence variants in South-East
RT Asia.";
RL J. Gen. Virol. 76:211-215(1995).
DR EMBL; U14256; AAC53945.1; -, Genomic RNA.
DR HSP; P27958; IHEI.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.

```

FT NON TER 1 1
SQ SEQUENCE 138 AA; 15085 MW; D80B1B4880A961B1 CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 2; Length 138;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 70 PDKEVLY 76

RESULT 399
Q5J6J6_MYCVN
ID Q5J6J6_MYCVN PRELIMINARY; PRT; 142 AA.
AC Q5J6J6;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DE Hypothetical protein.
OS Mycobacterium vanbaalenii.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=110539;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PYR-1;
RA Brezna B., Khan A.A., Freeman J.P., Stingley R.L., Cerniglia C.E.;
RT "Role of two cytochromes p450 in polycyclic aromatic hydrocarbons
RL degradation by Mycobacterium vanbaalenii PYR-1."
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY496703; AAS77251.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 142 AA; 15131 MW; 218A7B37A5281A6D CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 2; Length 142;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 7 VLAALAA 13

RESULT 400
Q99KF4_MOUSE
ID Q99KF4_MOUSE PRELIMINARY; PRT; 146 AA.
AC Q99KF4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Slec9a1 protein (Fragment).
GN Name=Slec9a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II;
RA Tissue=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Klausner R.D., Collins F.S., Wagner L., Shemen L., Schuler G.D.,
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Locuallano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```

```

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettman M., Madao A., Rodrigues S., Sanchez A.,
RA Whiting M., Madao A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CZECH II;
RA Tissue=Mammary tumor metastasized to lung. Tumor arose spontaneously;
RL Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004687; AA04687.1; -; mRNA.
DR Ensembl; ENSMUSG0000028854; Mus musculus.
DR MGI; MGI:102462; Slc9a1.
DR GO; GO:0016323; C:basolateral plasma membrane; TAS.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005624; C:membrane fraction; IDA.
DR GO; GO:0015385; F:sodium:hydrogen antiporter activity; IDA.
DR GO; GO:0006885; P:regulation of pH; IDA.
DR GO; GO:0006814; P:sodium ion transport; IDA.
DR InterPro; IPR001970; NaH exchngr_1.
DR PRINTS; PR01085; NAHEXCHGR1.
FT NON TER 1
SQ SEQUENCE 146 AA; 15831 MW; 2EB9EB778DF7B9C CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 2; Length 146;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGVVLGL 89
DB 68 KGVVLGL 74

RESULT 401
Q8TW80_METKA
ID Q8TW80_METKA PRELIMINARY; PRT; 147 AA.
AC Q8TW80;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Uncharacterized protein specific for M.kandleri, MK-24 family.
GN OrderedLocustNames=MK1156;
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014; DOI=10.1073/pnas.032671499;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozhavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AB010406; AAM02369.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 147 AA; 15719 MW; 63CC16D762BEE5FF CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 2; Length 147;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 20 VLAALAA 26
Db 7 VLAALAA 13

RESULT 402
Q521A9 MAGGR PRELIMINARY; PRT; 151 AA.
AC Q521A9
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=MG05191.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaportha.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachanteang P., Baldwin J., Barry A.,
RA Bayul T., Blitshetev B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowaky M., Boukhaltel B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gairke A., Goyette A., Graham J., Grandbois E., Gyalteen K., Hafez N.,
RA Hegopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kanat A., Kamvesselis M., Karlsson E.,
RA Kells C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., MacDonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Masirov J., Mihalev L., Mihova T., Mikkelson T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'neill K., Oseman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnuez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone S., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Zimmer A., Zody M., Zander E.;
RT "The genome sequence of Magnaporthe grisea."
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RA Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC preliminary data.
DR EMBL: AACU01000795; EAS52499.1; -; Genomic_DNA.
SQ SEQUENCE 151 AA; 16356 MW; 931F73E2098AE85 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVUL 21
Db 56 VLLGGVUL 62

RESULT 403
Q72BY0 DESVH PRELIMINARY; PRT; 155 AA.
AC Q72BY0;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Terminase, small subunit.
GN OrderedLocuNames=DVU1504;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.P., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., Deboy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidssen T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough."
RL Nat. Biotechnol. 22:554-559(2004).
RA EMBL: AB017314; AAS95982.1; -; Genomic_DNA.
DR TIGR; DVU1504; -
DR InterPro; IPR006448; Sm term P27.
DR Pfam; PF05119; Terminase_4; 1.
DR TIGRfams; TIGR01558; sm_term_P27; 1.
DR Complete proteome.
SQ SEQUENCE 155 AA; 17012 MW; B83008930170BC73 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAYC 28
Db 68 AALAAAYC 74

RESULT 404
Q57VK1 9TRYP PRELIMINARY; PRT; 159 AA.
AC Q57VK1;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=Tb927.7.6100;
OS Trypanosoma brucei
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,

```



```

RA Shallom J., Hou L., Djikeng A., Feldblyum T., Hostettler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Fai G., Peterson J.,
RA Khalak H.G., Salasberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC159439; AAX70368.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 159 AA; 17877 MW; 9337230DC8363718 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLA 22
DB 23 LGGVLA 29
|||||

RESULT 405
Q92WY2 RHIME PRELIMINARY; PRT; 159 AA.
AC Q92WY2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein SMD20180.
GN OrderedLocNames=RB0180; ORFNames=SMB20180;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=1021;
EX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL591985; CAC48580.1; -; Genomic_DNA.
DR PIR; D95864; D95864.
KW Complete proteome; Hypothetical protein; Plasmid.
SQ SEQUENCE 159 AA; 16605 MW; FAB997995102EFD5 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 159;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 17 VLAALAA 23
|||||

RESULT 406
Q7YG75_9GOBI PRELIMINARY; PRT; 160 AA.
AC Q7YG75;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 2 (Fragment).
GN Name=ND2;
OS Gnatholepis anjerensis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Perciformes; Gobioidae;
OC Gobiidae; Gnatholepis.
OX NCBI_TaxID=194857;
[1]

```

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=142;
RA Thacker C.E.;
RT "Phylogeny and species boundaries in the gobiid genus Gnatholepis
RT (Teleostei: Perciformes).";
RL Zool. J. Linn. Soc. 142:573-582(2004).
DR EMBL; AF537733; AAP46814.1; -; Genomic_DNA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR GO; GO:0006120; P:mitochondrial electron transport, NADH to u. . .; IEA.
DR InterPro; IPR010933; NADH dehy_S2 C.
DR Pfam; PF06444; NADH dehy_S2 C; 1.
DR PRINTS; PR01436; NADHDHGNASE2.
KW Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 160 AA; 17255 MW; 0D8BD25962E8F631 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 160;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 91 VLAALAA 97
|||||

RESULT 407
Q4NR09_9DELT PRELIMINARY; PRT; 161 AA.
AC Q4NR09;
DT 13-SEP-2005 (TREMBlrel. 31, Created)
DT 13-SEP-2005 (TREMBlrel. 31, Last sequence update)
DT 13-SEP-2005 (TREMBlrel. 31, Last annotation update)
DE Hypothetical protein precursor.
GN ORFNames=AdelDRAFT_1108;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacterales; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Iserani S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RA Laximer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000030; EAL78039.1; -; Genomic_DNA.
KW Hypothetical protein; Signal.
FT SIGNAL 1 30
SQ SEQUENCE 161 AA; 16348 MW; 4905D0382CB497E CRC64;

Query Match          5.9%; Score 7; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||

```

Db 17 VLAALAA 23

RESULT 408
 Q514C4_PECGU
 ID Q514C4_PECGU PRELIMINARY; PRT; 168 AA.
 AC Q514C4;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 OS Pectinaria gouldii (Trumpet worm) (Ice-cream cone worm).
 OC Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Canalipalpata;
 OC Terebellida; Pectinariidae; Pectinaria.
 OX NCBI_TaxID=260746;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Van Dyke A.J., Tauer T.J.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 RM EMBL; AY856416; AA047576.1; -, mRNA.
 KW Hypothetical protein.
 SQ SEQUENCE 168 AA; 17973 MW; C4B462082DC4493 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 168;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 13 VLAALAA 19

RESULT 409
 Q9HQ7_HALSA
 ID Q9HQ7_HALSA PRELIMINARY; PRT; 172 AA.
 AC Q9HQ7;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Vng1036h.
 OS OrderedLocusNames=VNG1036H;
 GN Halobacterium salinarum (Halobacterium halobium).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=2242;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
 RX NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlachroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL; AE005037; AAG19446.1; -, Genomic_DNA.
 DR PIR; B84260;
 DR InterPro; IPR006311; Tst.
 DR TIGRPFAMS; TIGR01409; TAT_signal_seq; 1.
 KW Complete proteome.
 SQ SEQUENCE 172 AA; 18002 MW; C466A676D0365BA6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 8 VLAALAA 14

RESULT 410
 Y433_AERPE
 ID Y433_AERPE STANDARD; PRT; 174 AA.
 AC Q9YF05;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Hypothetical UPF0290 protein APE0433.
 GN OrderedLocusNames=APE0433;
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KL;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Negai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix KL.";
 RL DNA Res. 6:83-101(1999).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: Belongs to the UPF0290 family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; BA000002; BAA79391.1; -, Genomic_DNA.
 DR PIR; C72737; C72737.
 DR HAMAP; MF 01117; -, 1.
 DR InterPro; IPR002726; DUF46.
 DR Pfam; PF01864; DUF46; 1.
 KW Complete proteome; Hypothetical protein; Transmembrane.
 FT TRANSMEM 15 37 Potential.
 FT TRANSMEM 58 80 Potential.
 FT TRANSMEM 84 106 Potential.
 FT TRANSMEM 126 143 Potential.
 FT TRANSMEM 147 169 Potential.
 SQ SEQUENCE 174 AA; 17953 MW; BD4C321CB0979DBE CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 174;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
 |||||
 Db 73 GVLAALA 79

RESULT 411
 Q5P936_AZOSE
 ID Q5P936_AZOSE PRELIMINARY; PRT; 177 AA.
 AC Q5P936;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein.
 OS OrderedLocusNames=AZOSEA00530; ORFNames=eba93;
 OS Azococcus sp. (strain EDN1).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
 OC Rhodocyclaceae; Azococcus.
 OX NCBI_TaxID=76114;
 RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=EBN1;
RX PubMed=1551059; DOI=10.1007/s00203-004-0742-9;
RA Rabus R., Kube M., Heider J., Beck A., Heitmann K., Widdel F.,
RA Reinhardt R.;
RT "The genome sequence of an anaerobic aromatic-degrading denitrifying
RT bacterium, strain EBN1.";
RL Arch. Microbiol. 183:27-36(2005).
DR EMBL; CR55306; CA106173.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 177 AA; 19966 MW; C32AACB2E9D576BA CRC64;

Query Match          5.9%; Score 7; DB 2; Length 177;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      16 LLGGVLA 22
      |||||
Db      108 LLGGVLA 114

RESULT 412
Q67LJ9 SYMTH
ID Q67LJ9 SYMTH PRELIMINARY; PRT; 178 AA.
AC Q67LJ9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
DE OrderedLocustNames=STH2462;
OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IAM14863;
RX PubMed=15383646; DOI=10.1093/nar/gkh830;
RA Ueda K., Yamashita A., Ishikawa J., Shinada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Genome sequence of Symbiobacterium thermophilum, an uncultivable
RT bacterium that depends on microbial commensalism.";
RL Nucleic Acids Res. 32:4937-4944(2004).
DR EMBL; AF06840; BAD1447.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 178 AA; 19505 MW; 5F5E5BA25C10DB6A CRC64;

Query Match          5.9%; Score 7; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      20 VLAALAA 26
      |||||
Db      31 VLAALAA 37

RESULT 413
FAIM1 MOUSE
ID FAIM1 MOUSE STANDARD; PRT; 179 AA.
AC Q9WUD8; Q9DAK6;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Fas apoptotic inhibitory molecule 1.
GN Name=Paim; Synonyms=Paim1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Muridae; Mus.
OX NCBI_TaxID=10090;
[1]
RN NUCLEOTIDE SEQUENCE, FUNCTION, AND TISSUE SPECIFICITY.
RC TISSUE=Thymus;
RX MEDLINE=99177249; PubMed=10075978; DOI=10.1084/jem.189.6.949;

RA Schneider T.J., Fischer G.M., Donohoe T.J., Colarusso T.P.,
RA Rothstein T.L.;
RT "A novel gene coding for a Fas apoptosis inhibitory molecule (FAIM)
RT isolated from inducibly Fas-resistant B lymphocytes.";
RL J. Exp. Med. 189:949-956(1999).
[2]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Badarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Giasi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Khan A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[3]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RP STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Plays a role as an inducible effector molecule that
CC mediates Fas resistance produced by surface Ig engagement in B
CC cells.
CC -!- TISSUE SPECIFICITY: Widely expressed, with the highest levels in
CC brain, thymus, kidney, and spleen.
CC -!- SIMILARITY: Belongs to the FAIM1 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not

```

```
CC removed.
CC -----
DR EMBL; AF130367; AA023879.1; -; mRNA.
DR EMBL; AK005762; BAB24225.1; -; mRNA.
DR EMBL; BC079662; AAH79662.1; -; mRNA.
DR EMBL; ENSMUSG00000032463; Mus musculus.
DR MGI; MGI:1344387; Faim.
DR GO; GO:0006916; P:anti-apoptosis; IDA.
DR InterPro; IPR010695; FAIM.
DR PANTHER; PTHR13088; FAIM; 1.
DR Pfam; PF06905; FAIM; 1.
KW Apoptosis.
FT CONFLICT 157 157 G -> R (in Ref. 2).
SQ SEQUENCE 179 AA; 20201 MW; 10849A8B65926556 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 TTSTWVL 16
Db 96 TTSTWVL 102

RESULT 414
Q8ECD8 SHEON
ID Q8ECD8_SHEON PRELIMINARY; PRT; 179 AA.
AC Q8ECD8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein S03205.
GN OrderedLocNames=S03205;
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Shewanellaceae; Shewanella.
OC NCBI_TaxID=70863;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Unayam L.A.,
RA White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA Venter J.C., Nelson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015758; AAN56204.1; -; Genomic_DNA.
DR TIGR; S03205; -.
KW Complete proteome.
SQ SEQUENCE 179 AA; 19376 MW; ECE90BB932473EDD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 EQAQVIA 79
Db 59 EQAQVIA 65

RESULT 415
Q84XX9 BRAP
ID Q84XX9_BRAP PRELIMINARY; PRT; 184 AA.
AC Q84XX9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
```

```
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HUA2 (Fragment).
OS Brassica rapa subsp. pekinensis (Chinese cabbage) (Celery cabbage).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=15351;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Leaf;
RA Jang H., Hur Y.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176678; AA022522.1; -; mRNA.
FT NON_TER 1 184
FT NON_TER 184 184
SQ SEQUENCE 184 AA; 19709 MW; 7CA4CD7F183B34BD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIVP 50
Db 20 GKPAIVP 26

RESULT 416
Q73VL0 MYCPA
ID Q73VL0_MYCPA PRELIMINARY; PRT; 184 AA.
AC Q73VL0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=MAP3003C;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017238; AAS05551.1; -; Genomic_DNA.
DR InterPro; IPR004398; Cons hypoth95.
DR InterPro; IPR002052; N6 Mcase.
DR Pfam; PF03602; Cons hypoth95; 1.
DR TIGRFAMs; TIGR00095; Cons hypoth95; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 184 AA; 18870 MW; BF50B09C0079B204 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 132 VLAALAA 138

RESULT 417
Q7U864 SYNFX
ID Q7U864_SYNFX PRELIMINARY; PRT; 185 AA.
AC Q7U864;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SYNM0760;
OS Synechococcus sp. (strain WH8102).
```

```

OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=84588;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
RA Paulsen I.T., Duirene A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synechococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL: BX569691; CAE07275.1; -: Genomic_DNA.
DR InterPro: IPR002610; Rhomboid_like.
DR Pfam: PF01694; Rhomboid; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 185 AA; 19709 MW; 8DRCF1A6AEF56A6D CRC64;

Query Match 5.9%; Score 7; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
Db 163 LGGVLA 169

RESULT 418
Q68QF2 NPVAP
ID Q68QF2 NPVAP PRELIMINARY; PRT; 185 AA.
AC Q68QF2
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Fibroblast growth factor.
GN Name=fgf;
OS Antheraea pernyi nuclear polyhedrosis virus (AnpV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
CC Nucleopolyhedrovirus.
OX NCBI_TaxID=161494;
RN NUCLEOTIDE SEQUENCE.
RA Song Z., Wang W., Shen W., Wang L.;
RT "Cloning and Analysis of Fibroblast Growth Factor Gene of Antheraea
RT pernyi Nuclear Polyhedrosis Virus.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY641527; AAU00991.1; -: Genomic DNA.
DR GO: GO:0008083; F: growth factor activity; IEA.
DR GO: GO:0003700; E: transcription factor activity; IEA.
DR GO: GO:0006355; P: regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR002197; HTH Fis.
DR Pfam: PF00167; FGF; 1.
DR PRINTS: PR01590; HTHFIS.
DR ProDom: PD000831; IL1_HBGF; 1.
DR SMART: SM00442; FGF; 1.
SQ SEQUENCE 185 AA; 21097 MW; 0AD1C991DD5BCCA5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 185;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 5 VLAALAA 11

RESULT 419
Q91XV9_9HEPC
ID Q91XV9_9HEPC PRELIMINARY; PRT; 186 AA.
AC Q91XV9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Polyprotein (Fragment).

```

```

OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=20230065; PubMed=10764648; DOI=10.1126/science.288.5464.339;
RA Farci P., Shioda A., Coiana A., Diaz G., Peddis G., Meppolder J.C.,
RA Strazzer A., Chien D.X., Munoz S.J., Balestrieri A., Purcell R.H.,
RA Alter H.J.;
RT "The outcome of acute hepatitis C predicted by the evolution of the
RT viral quasispecies.";
RL Science 288:339-344(2000).
DR EMBL: AF246089; AAF66388.1; -: Genomic RNA.
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0019031; C: viral envelope; IEA.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV_NSI.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV_NSI; 1.
KW Envelope protein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 186
FT NON_TER 186
SQ SEQUENCE 186 AA; 20145 MW; 3A05A6D328013326 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 37 GVLAALA 43

RESULT 420
Q8QRV4_9BETA
ID Q8QRV4_9BETA PRELIMINARY; PRT; 187 AA.
AC Q8QRV4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Immediate-early glycoprotein US3.
OS Pongine herpesvirus 4 (Chimpanzee cytomegalovirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=188763;
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22421467; PubMed=12533697; DOI=10.1099/vir.0.18606-0;
RA Davison A.J., Dolan A., Akter P., Addison C., Dargan D.J.,
RA Alcendor D.J., McGeoch D.J., Hayward G.S.;
RT "The human cytomegalovirus genome revisited: comparison with the
RT chimpanzee cytomegalovirus genome.";
RL J. Gen. Virol. 84:17-28(2003).
DR EMBL: AF480884; AAM00784.1; -: Genomic_DNA.
DR InterPro: IPR009237; Cytomegalovirus US3.
DR Pfam: PF05963; Cytomegalovirus US3; 1.
SQ SEQUENCE 187 AA; 21636 MW; 24A8075A60C5E66C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 6 VLAALAA 12

RESULT 421
Q64D23_9ARCH
ID Q64D23_9ARCH PRELIMINARY; PRT; 188 AA.
AC Q64D23;
DT 25-OCT-2004 (TREMBlrel. 28, Created)

```

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Sulfopyruvate decarboxylase alpha chain (Fragment).
GN ORFNames=GZ19A5_46;
OS uncultured archaeon GZfosi19A5.
OC Archaea; environmental samples.
OX NCBI_TaxID=285362;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15353801; DOI=10.1126/science.1100025;
RA Hallam S.J., Putnam N., Preston C.M., Dettler J.C., Rokhsar D.,
RA Richardson P.M., DeLong E.F.;
RT "Reverse methanogenesis: testing the hypothesis with environmental
RT genomics";
RL Science 305:1457-1462(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Putnam N., Dettler J.C., Richardson P.M., Rokhsar D.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY714829; AUA82704.1; -, Genomic_DNA.
KW Pyruvate.
FT NON TER 188 188
SQ SEQUENCE 188 AA; 21281 MW; E4B0E1976E5645B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GGRPAIV 49
Db 50 GGRPAIV 56
|||||
- - - - -

RESULT 422
Q57D47 BRUAB PRELIMINARY; PRT; 188 AA.
ID Q57D47;
AC Q57D47;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Transcriptional regulator, TetR family.
GN OrderedLocuNames=Brub1_1096;
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=9-941 / Biovar 1;
RX PubMed=15805518; DOI=10.1128/JB.187.8.2715-2726.2005;
RA Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuerner R.L.,
RA Qiang Z., Li L.-L., Kapur V., Alt D.P., Olsen S.C.;
RT "Completion of the genome sequence of Brucella abortus and comparison
RT to the highly similar genomes of Brucella melitensis and Brucella
RT suis.";
RL J. Bacteriol. 187:2715-2726(2005).
CC -1- FUNCTION: Repressor involved in choline regulation of the bet
CC genes (By similarity).
CC -1- PATHWAY: Betaine biosynthesis from choline; regulation.
CC EMBL; AE017223; AAX74437.1; -, Genomic DNA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTH_TETR.
DR PROSITE; PS0977; HTH_TETR 2; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 188 AA; 20679 MW; F7F0102B855225AD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 105 GVLAALA 111
|||||
- - - - -

RESULT 423
Q8G0K2 BRUSU
ID Q8G0K2 BRUSU PRELIMINARY; PRT; 188 AA.
AC Q8G0K2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcriptional regulator, TetR family.
GN OrderedLocuNames=BR1090;
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Unayam L.A., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153(2002).
CC -1- FUNCTION: Repressor involved in choline regulation of the bet
CC genes (By similarity).
CC -1- PATHWAY: Betaine biosynthesis from choline; regulation.
CC EMBL; AE014291; AAN30010.1; -, Genomic_DNA.
DR PIR; AR3363; AR3363.
DR TIGR; BR1090; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTH_TETR.
DR PROSITE; PS0977; HTH_TETR 2; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 188 AA; 20679 MW; B6F0102AD45104AA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 105 GVLAALA 111
|||||
- - - - -

RESULT 424
Q8YHBI BRUME
ID Q8YHBI BRUME PRELIMINARY; PRT; 188 AA.
AC Q8YHBI;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE TRANSCRIPTIONAL REGULATOR, TETR FAMILY.
GN OrderedLocuNames=BWE10891;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;

```

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=21642680; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mujer C., Los T.,
RA Ivanova N., Anderson N., Bhatnagar V.A., Lykidis A., Reznik G.,
RA Jablonka L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kypides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
  Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -!- FUNCTION: Repressor involved in choline regulation of the bet
  genes (By similarity).
CC -!- PATHWAY: Betaine biosynthesis from choline; regulation.
DR EMBL; AB009529; AAL52072.1; -; Genomic_DNA.
DR PIR; AB3363; AB3363.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR012487; Homeodomain-rel.
DR InterPro; IPR001647; HTH Tetr.
DR Pfam; PF00440; Tetr_N; 1.
DR PRINTS; PR00455; HTHTETR.
DR PROSITE; PS0977; HTH TETR 2; 1.
KW Complete proteome, DNA-binding; Transcription;
  Transcription regulation.
SQ SEQUENCE 188 AA; 20679 MW; B6F0102AD45104AA CRC64;

Query Match          5.9%; Score 7; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAA 25
DB 105 GVLAA 111
|||||
|||||

RESULT 425
GCH11_PSEPK
ID GCH11_PSEPK STANDARD; PRT; 190 AA.
AC Q88LV4;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE GTP cyclohydrolase I 1 (EC 3.5.4.16) (GTP-CH-I 1).
GN Name=folE1; OrderedLocusNames=PP1823;
OS Pseudomonas putida (strain KT2440);
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RX MEDLINE=22421060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.C., DeBoy R.T., Daugherty S.C., Kolonay J.P.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Morzez A., Uytterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
  metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
CC -!- CATALYTIC ACTIVITY: GTP + 2 H(2)O = formate + 2-amino-4-hydroxy-6-
  (erythro-1,2,3-trihydroxypropyl)-dihydropteridine triphosphate.
CC -!- PATHWAY: Cofactor biosynthesis; tetrahydrofolate biosynthesis; 2-
  amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine diphosphate
  from GTP: step 1.
CC -!- SUBUNIT: Homopolymer (By similarity).
CC -!- SIMILARITY: Belongs to the GTP cyclohydrolase I family.

```

```

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
DR EMBL; AB016780; AAN67442.1; -; Genomic_DNA.
DR HSSP; P22288; 11S8.
DR TIGR; PP1823; -.
DR HAMAP; MF 00223; -; 1.
DR InterPro; IPR001474; GTP_cyclohydrol.
DR Pfam; PF01227; GTP_cyclohydrol; 1.
DR ProDom; PD003330; GTP_cyclohydrol; 1.
DR TIGRFAMs; TIGR00063; folE; 1.
DR PROSITE; PS00859; GTP_CYCLOHYDROL_1_1; 1.
DR PROSITE; PS00860; GTP_CYCLOHYDROL_1_2; 1.
KW Complete proteome; Hydrolase; One-carbon metabolism.
FT DISULFID 82 153 By similarity.
SQ SEQUENCE 190 AA; 21388 MW; E65550A866122653 CRC64;

Query Match          5.9%; Score 7; DB 1; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGVVLGL 89
DB 100 KGVVLGL 106
|||||
|||||

RESULT 426
OSQMN8_ORYSA
ID OSQMN8_ORYSA PRELIMINARY; PRT; 192 AA.
AC OSQMN8;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein P0419B01.2.
GN Name=P0419B01.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RX NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohka I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn H.H., Kim H.I., Run M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
DR EMBL; AP003244; BAD73317.1; -; Genomic_DNA.
DR InterPro; IPR004895; PRAL.
DR Pfam; PF03208; PRAL; 1.
KW Hypothetical protein.
SQ SEQUENCE 192 AA; 20202 MW; 2AC22BB141B3647E CRC64;

Query Match          5.9%; Score 7; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26

```

```

Db      139 VLALAA 145
|||||
RESULT 427
Q9LCT1 BRAJA PRELIMINARY; PRT; 192 AA.
AC Q9LCT1_079V12;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (blr0682 protein).
GN OrderedLocuNames=blr0682;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=110spc4;
RX MEDLINE=97261868; PubMed=9108282; DOI=10.1007/s004380050408;
RA Minder A.C., Narberhaus F., Babst M., Hennecke H., Fischer H.-M.;
RT "The dnaJ operon belongs to the sigma32-dependent class of heat shock
RT genes in Bradyrhizobium japonicum.";
RL Mol. Gen. Genet. 254:195-206(1997).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=110spc4;
RX MEDLINE=20082837; PubMed=10613857;
RA Minder A.C., Fischer H.-M., Hennecke H., Narberhaus F.;
RT "Role of HrcA and CIRCE in the heat shock regulatory network of
RT Bradyrhizobium japonicum.";
RL J. Bacteriol. 182:14-22(2000).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=110spc4;
RX MEDLINE=21150465; PubMed=11251836;
RA Minder A.C., de Rudder K.E.E., Narberhaus F., Fischer H.-M.,
RA Hennecke H., Geiger O.;
RT "Phosphatidylcholine levels in Bradyrhizobium japonicum membranes are
RT critical for an efficient symbiosis with the soybean host plant.";
RL Mol. Microbiol. 39:1186-1198(2001).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=110spc4;
RA Narberhaus F.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBSJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamieawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iziguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; Y09633; CAB91879.1; -; Genomic DNA.
DR EMBL; BA000040; BAC45947.1; -; Genomic_DNA.
DR InterPro; IPR005025; FNM_red.
DR Pfam; PF03358; FNM_red; 1.
SQ Complete proteome; Hypothetical protein.
KW SEQUENCE 192 AA; 20669 MW; 7CE30F8A200511B7 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      21 LAALAA 27
|||||
Db      22 LAALAA 28
|||||

RESULT 428
Q7NPH4 GLOVI PRELIMINARY; PRT; 197 AA.
AC Q7NPH4;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glr0081 protein.
GN OrderedLocuNames=gir0081;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacterales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tauchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsumoto A., Nakazaki N., Shimpo S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; BA000045; BAC88022.1; -; Genomic_DNA.
DR InterPro; IPR002610; Rhomboid_like.
DR Pfam; PF01694; Rhomboid; 1.
KW Complete proteome.
SQ SEQUENCE 197 AA; 21654 MW; BP5974F1A26F5BA5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      17 LGGVLA 23
|||||
Db      181 LGGVLA 187
|||||

RESULT 429
VP10 BPBRD STANDARD; PRT; 203 AA.
AC P28732;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Protein P10.
GN Name=X;
OS Bacteriophage PRD1.
OC Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.
OX NCBI_TaxID=10658;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=91306449; PubMed=1853567;
RA Bamford J.K.H., Haenninen A.-J., Pakula T.M., Ojala P.M.,
RA Kalkkinen N., Frilander M., Bamford D.H.;
RT "Genome organization of membrane-containing bacteriophage PRD1.";
RL Virology 183:658-676(1991).
CC -1- FUNCTION: The major coat protein P3 and two assembly factors (P10
CC and P17) are needed during the assembly of the virus particle
CC inside the host cell, when the capsid protein multimers are
CC capable of enclosing the viral membrane from the host plasma
CC membrane containing the virus-encoded membrane-associated
CC proteins.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M69077; -; NOT_ANNOTATED_CDS; Genomic_DNA.
CC PIR; E36777; WMBPTB.

```


DR InterPro; IPR000637; AT_hook_DNA_bd.
 DR Pfam; PF02178; AT_hook; 2.
 KW Capsid assembly.
 SQ SEQUENCE 203 AA; 20688 MW; 064C2BB9C08151D9 CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 203;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 68 AAPYIEQ 74
 |||||
 DB 190 AAPYIEQ 196
 |||||
 RESULT 430
 ID Q6EDX4_BPPRD PRELIMINARY; PRT; 203 AA.
 AC Q6EDX4;
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE Assembly protein.
 OS Bacteriophage PR772.
 OC Viruses; dsDNA viruses, no RNA stage; Tectiviridae; Tectivirus.
 OX NCBI_TaxID=261665;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lute S.C., Aranha H., Tremblay D., Liang D., Ackermann H.-W., Chu B.,
 RA Moineau S., Brorson K.A.;
 RT "Characterization of the coliphage PR772 and evaluation of its use for
 RT virus filter performance testing."
 RL Appl. Environ. Microbiol. 70:4864-4871(2004).
 DR EMBL; AY441783; AAR9750.1; -; Genomic DNA.
 SQ SEQUENCE 203 AA; 20687 MW; 29DD1143DE5F9F31 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 203;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 68 AAPYIEQ 74
 |||||
 DB 190 AAPYIEQ 196
 |||||
 RESULT 431
 ID Q6N9H6_RHOPA PRELIMINARY; PRT; 204 AA.
 AC Q6N9H6;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 DE LemA family precursor.
 OS OrderedLocusNames=RPA1573;
 GN Rhodopseudomonas palustris.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Rhodospseudomonas.
 OX NCBI_TaxID=1076;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CGA009 / ATCC BAA-98;
 RX PubMed=14704707; DOI=10.1038/nbt923;
 RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
 RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
 RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
 RA Harrison F.H., Gibson J., Harwood C.S.;
 RT "Complete genome sequence of the metabolically versatile
 RT photosynthetic bacterium Rhodopseudomonas palustris."
 RL Nat. Biotechnol. 22:55-61(2004).
 DR EMBL; BX572598; CAE27014.1; -; Genomic_DNA.
 DR InterPro; IPR007156; LemA.
 DR Pfam; PF04011; LemA; 1.
 KW Complete proteome; Signal.
 FT SIGNAL 1 21 Potential.

SQ SEQUENCE 204 AA; 22473 MW; 309584C261B134B8 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 204;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLAALAA 26
 |||||
 DB 7 VLAALAA 13
 |||||
 RESULT 432
 ID Q930S6_RHIME PRELIMINARY; PRT; 204 AA.
 AC Q930S6;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Putative transcriptional regulator.
 GN OrderedLocusNames=RA0119; ORFNames=SWa0223;
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OG Plasmid pSymA.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=1021;
 RX MEDLINE=21396509; PubMed=1481432; DOI=10.1073/pnas.161294798;
 RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
 RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
 RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
 RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
 RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
 RT "Nucleotide sequence and predicted functions of the entire
 RT Sinorhizobium meliloti pSymA megaplasmid."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
 CC -!- FUNCTION: Repressor involved in choline regulation of the bet
 CC genes (By similarity).
 CC -!- PATHWAY: Betaine biosynthesis from choline; regulation.
 DR EMBL; AB007206; AAK64777.1; -; Genomic_DNA.
 DR FIR; G95276; G95276.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR012287; Homeodomain-rel.
 DR InterPro; IPR001647; HTH_Tetr.
 DR Pfam; PF00440; Tetr_N; 1.
 DR PRINTS; PR00455; HTHTEPR.
 DR PROSITE; PS0977; HTH_TETR 2; 1.
 KW Complete proteome; DNA-binding; Plasmid; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 204 AA; 22388 MW; 6D786169A69304D0 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 204;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 GVLAALA 25
 |||||
 DB 119 GVLAALA 125
 |||||
 RESULT 433
 ID Q5BZH4_SCHJA PRELIMINARY; PRT; 206 AA.
 AC Q5BZH4;
 DT 10-MAY-2005 (TRENBLrel. 30, Created)
 DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
 OC Schistosomatoidea; Schistosomatidae; Schistosoma.

```
OX NCBI_TaxID=6182;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Han Z.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
RE EMBL; AY811312; AAX27201.1; -, mRNA.
DR InterPro; IPR010980; Cyt_c_b562.
KW Hypothetical protein.
SQ SEQUENCE 206 AA; 21665 MW; 0797165252DC8B01 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 206;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LEVTTST 13
Db 39 LEVTTST 45
|||||

RESULT 434
Q4WFS9_ASPFU
ID Q4WFS9_ASPFU PRELIMINARY; PRT; 215 AA.
AC Q4WFS9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE GPI anchored cell wall protein, putative.
GN ORFNames=AFU301150;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AF293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H.Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman N., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fooker M., Fraser A., Garcia J.L., Garcia M.J., Goble A.,
RA Goldman G.H., Gomi K., Griffith-Jones S., Gwilliam R., Haas B.,
RA Haas H., Harris D., Horiuchi H., Huang J., Humphrey S., Jimenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafont A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majors W.H., May G.S., Miller B.L., Mahmoud Y., Molina M., Monod M.,
RA Mouyna I., Mulligan S., Murphy L., O'Neil S., Paulsen I.,
RA Panalva M.A., Perte M., Price C., Pritchard B.L., Quail M.A.,
RA Rabinowitsch E., Rawlins N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,
RA White O., Woodward J., Yu J.-H., Fraser C., Galagan J.E., Asai K.,
RA Machida M., Hall N., Barrell B., Denning D.W.;
RT "Genomic sequence of the pathogenic and allergenic filamentous fungus
Aspergillus fumigatus";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAF01000010; EAL86398.1; -; Genomic DNA.
DR EMBL; AAF01000010; EAL86398.1; -; Genomic DNA.
SQ SEQUENCE 215 AA; 21591 MW; AFB00AF23ED8ADD8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 7 VLAALAA 13
|||||

RESULT 435
Q9J7E1_NEIG1
ID Q9J7E1_NEIG1 PRELIMINARY; PRT; 217 AA.
AC Q9J7E1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=NGO1237;
OS Neisseria gonorrhoeae (strain ATCC 700825 / FA 1090).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=242231;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA Lewis L.A., Gillaspay A.F., McLaughlin R.E., Gipson M., Ducey T.F.,
RA Ownbey T., Hartman K., Nydick C., Carson M.B., Vaughn J., Thomson C.,
RA Song L., Lin S., Yuan X., Najjar F., Zhan M., Ren Q., Zhu H., Qi S.,
RA Kanton S.M., Lai H., White J.D., Clifton S., Roe B.A., Dyer D.W.;
RT "The complete genome sequence of Neisseria gonorrhoeae";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE004969; AAW89896.1; -; Genomic DNA.
DR InterPro; IPR003782; SCO1_SenC; 1.
DR Pfam; PF02630; SCO1_SenC; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 217 AA; 23107 MW; C6623B3BBE2247B7 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 14 VLAALAA 20
|||||

RESULT 436
Q9J7I0_NEIMA
ID Q9J7I0_NEIMA PRELIMINARY; PRT; 217 AA.
AC Q9J7I0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocuNames=NMA1767;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RC MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RX Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491";
RL Nature 404:502-506(2000).
DR EMBL; AL162757; CAB84995.1; -; Genomic DNA.
DR PIR; G81801; G81801.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003782; SCO1_SenC.
DR PANTHER; PTHR12151; SCO1_SenC; 1.
DR Pfam; PF02630; SCO1_SenC; 1.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 217 AA; 23268 MW; 862989950A147AAC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy 20 VLAALAA 26
Db 14 VLAALAA 20

RESULT 437
Q9JYH9 NEIMB
ID Q9JYH9_NEIMB PRELIMINARY; PRT; 217 AA.
AC Q9JYH9;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=NM1578;
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OC NCBI_TaxID=491;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=2017555; PubMed=10710307; DOI=10.1126/science.287.5459.1809;
RA Tetelin H., Saunders N.J., Heidelberg J.F., Jeffries A.C.,
RA Nelson K.E., Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F.,
RA Dodson R.J., Nelson W.C., Gwinn M.L., DeBoy R.T., Peterson J.D.,
RA Hickey E.K., Haft D.H., Salzberg S.L., White O., Fleischmann R.D.,
RA Dougherty B.A., Mason T.M., Ciecko A., Parksey D.S., Blair E.,
RA Cittone H., Clark E.B., Cotton M.D., Utterback T.R., Khouri H.M.,
RA Qin H., Vamathevan J.J., Gill J., Scarlato V., Maignani V., Pizza M.,
RA Grandi G., Sun L., Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R.,
RA Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002098; AAP41931.1; -; Genomic_DNA.
DR PIR; B81067; B81067.
DR TIGR; NMB1578; -
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003782; SC01_SenC.
DR PANTHER; PTHR12151; SC01_SenC; 1.
DR Pfam; PF02630; SC01_SenC; 1.
DR Complete proteome; Hypothetical protein.
SQ SEQUENCE 217 AA; 23210 MW; 84PFB5997D8C47AAC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 14 VLAALAA 20

RESULT 438
Q9V4H2 DROME
ID Q9V4H2_DROME PRELIMINARY; PRT; 219 AA.
AC Q9V4H2; OSZB5;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 10-MAY-2005 (TREMELrel. 30, Last annotation update)
DE CG12838-PA (RE08073p).
GN Name=Tp42EO; ORFNames=CG12838;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

```

```

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
RA Laško P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Swirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Swirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;

```

```
RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoekins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.;
RT "Drosophila melanogaster release 4 sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RP NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
[7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN-Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarino H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003842; AAP59300.2; -; Genomic_DNA.
DR EMBL; AY070986; AAL48608.1; -; mRNA.
DR Ensembl; CG12838; Drosophila melanogaster.
DR FlyBase; FBgn0033136; CG12838.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR InterPro; IPR000301; Transmem 4.
DR Pfam; PF00335; Tetraspannin; I.
DR PRINTS; PR00259; TMPOUR.
DR SQU SEQUENCE 219 AA; 24487 MW; 0FCCAE370BCB2FA2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 23 GVLAALA 29
|||||

RESULT 439
Q4ZVF9_PSESY
ID Q4ZVF9_PSESY PRELIMINARY; PRT; 219 AA.
AC Q4ZVF9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Regulatory protein, GntR.
GN ORFNames=Psyr 1816;
OS Pseudomonas syringae pv. syringae B728a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=205918;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG DOE Joint Genome Institute;
RA Chain P., Larimer F., DiBartolo G., Copeland A., Lykidis A., Trong S.,
RA Land M., Goltzman E., Thiel J., Malfatti S., Lapidus A., Dettler J.C.,
RA Nolan M., Richardson P.M., Kyrpides N.C., Ivanova N.;
RT "Comparison of two complete genome sequences of Pseudomonas syringae
pv. syringae B728a and pv. tomato DC3000.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2005).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RG Loper J.;
RA Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B728a;
RA Feil H., Feil W.S., Lindow S.E.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000075; AAV36863.1; -; Genomic_DNA.
```

```
DR InterPro; IPR000524; HTH_GntR.
DR Pfam; PF00392; GntR; 1.
DR SMART; SM00345; HTH_GNTR; 1.
DR PROSITE; PS00949; HTH_GNTR; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 219 AA; 24639 MW; 54B6B6A9622AAC41 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 212 VLAALAA 218
|||||

RESULT 440
Q8DNP7_STRR6
ID Q8DNP7_STRR6 PRELIMINARY; PRT; 222 AA.
AC Q8DNP7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aquaporin Z-water channel protein.
GN Name=aqp2; OrderedLocusNames=spr1604;
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=171101;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21439245; PubMed=11544234;
RX DOI=10.1128/JB.183.19.5709-5717.2001;
RA Hoskins J., Estrom W.E. Jr., Arnold J., Blaszcak L.C., Burgett S.,
RA DeHoff B.S., Estrem S.T., Fritz L., Fu D.-J., Fuller W., Geringer C.,
RA Gilmour R., Glass J.S., Khoja H., Kraft A.R., Lagace R.E.,
RA LeBlanc D.J., Lee L.N., Lefkowitz E.J., Lu J., Matsushina P.,
RA McAhren S.M., McHenry M., McLeaster K., Mundy C.W., Niclas T.I.,
RA Norris F.H., O'Gara M., Peery R.B., Robertson G.T., Rockey P.,
RA Sun P.-M., Winkler M.E., Yang Y., Young-Bellido M., Zhao G.,
RA Zook C.A., Baltz R.H., Jaskunas S.R., Rostock P.R. Jr., Skatrud P.L.,
RA Glass J.I.;
RT "Genome of the bacterium Streptococcus pneumoniae strain R6.";
RL J. Bacteriol. 183:5709-5717(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AE008527; AAL00407.1; -; Genomic_DNA.
DR FIR; B98072; B98072.
DR HSP; P29972; IH61.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR00783; MINTINSICP.
DR ProDom; PD000295; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 222 AA; 22453 MW; 5A5E03F5083462B3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 206 GGVLAAL 212
|||||

RESULT 441
Q97P66_STRPN
ID Q97P66_STRPN PRELIMINARY; PRT; 222 AA.
```

```

AC Q97P66;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aquaporin.
GN OrderedLocustNames=SPI778;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC BAA-334 / TIGR4;
RX MEDLINE=21357209; PubMed=11463916; DOI=10.1126/science.1061217;
RA Tetelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S.N., Heidelberg J.F., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E.K., Khouri H.M., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S.V., Dickinson T.,
RA Hickey E.K., Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae."
RL Science 293:498-506(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the MIP/aquaporin (TC 1.A.8) family.
DR EMBL; AB007471; AAK75851.1; -; Genomic_DNA.
DR PIR; B95207; B95207.
DR HSP; P29972; IHE1.
DR TIGR; SPI778; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0019867; C: outer membrane; IEA.
DR GO; GO:0015288; F: porin activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR000425; MIP.
DR Pfam; PF00230; MIP; 1.
DR PRINTS; PR00783; MINTRINSICP.
DR PRODOM; PD000295; MIP; 1.
DR PROSITE; PS00221; MIP; 1.
KW Complete proteome; Porin; Transmembrane; Transport.
SQ SEQUENCE 222 AA; 22487 MW; 0D34A352655EC5DE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 206 GGVLAAL 212

RESULT 442
ID Q8NS26 CORGL PRELIMINARY; PRT; 222 AA.
AC Q8NS26; Q6M672;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE ABC-type transporter, ATPase component (ABC-type transport system,
DE involved in lipoprotein release, ATPase component).
DE OrderedLocustNames=Cgl0856, cg0978;
GN Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

```

```

RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=22830012; PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
RA Burkovski A., Duesch N., Eggeling L., Eikmanns B.J., Gaigalat L.,
RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Puhler A.,
RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
RA Tauch A.;
RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
RT and its impact on the production of L-aspartate-derived amino acids
RT and vitamins."
RL J. Biotechnol. 104:5-25(2003).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; BA000036; BAB98249.1; -; Genomic_DNA.
DR EMBL; BX927150; CAF19562.1; -; Genomic_DNA.
DR HSP; Q58206; IF30.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005524; F: ATP binding; IEA.
DR GO; GO:0016887; F: ATPase activity; IEA.
DR GO; GO:0016787; F: hydrolase activity; IEA.
DR GO; GO:0000166; F: nucleotide binding; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00005; ABC_tran; 1.
DR PRODOM; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome; Lipoprotein; Membrane;
KW Nucleotide-binding; Transport.
SQ SEQUENCE 222 AA; 23461 MW; 37A8347ECC0A2A2B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 GLLQRAT 94
Db 58 GLLQRAT 64

RESULT 443
QSTB22 HUMAN PRELIMINARY; PRT; 224 AA.
ID QSTB22;
AC QSTB22;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Solute carrier family 9 (Sodium/hydrogen exchanger), isoform 1
DE (Antiporter, Na+/H+, amiloride sensitive) (Fragment).
GN Name=SLC9A1; ORFNames=RP4-633N17.1-005;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
[1]
RN NUCLEOTIDE SEQUENCE.
RA Hall R.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137860; CA122087.1; -; Genomic_DNA.
FT NON TER 224 224
SQ SEQUENCE 224 AA; 25031 MW; 50552400B882D8A3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLGL 89

```

Db 159 KGKVLGL 165

RESULT 444

Q6N4M2 RHOPA PRELIMINARY; PRT; 224 AA.
AC Q6N4M2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Protease precursor.
GN OrderedLocusNames=RPA3315;
OS Rhodopseudomonas palustris;
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CQA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Peltier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.B., Bobet C., Torres y Torres J.L., Peres C.,
RA Harrison P.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572603; CA28756.1; -; Genomic DNA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR008915; Peptidase M50.
DR Pfam; PF02163; Peptidase M50; 1.
KW Complete proteome; Protease; Signal.
FT SIGNAL 1 32
SQ SEQUENCE 224 AA; 23914 MW; B8E99F0BPE427EB7 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 224;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 110 VLAALAA 116

RESULT 445

Q8A6N6 BACTN PRELIMINARY; PRT; 226 AA.
AC Q8A6N6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BTL841;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016933; AA076948.1; -; Genomic DNA.
DR InterPro; IPR007395; DUF_Zn_Bind.
DR Pfam; PF04298; Zn_peptidase_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 226 AA; 24929 MW; D51104A1PD43E418 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 228;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

Query Match 5.9%; Score 7; DB 2; Length 226;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 TWVLGG 19

Db 129 TWVLGG 135

RESULT 446

ISPD BACHD STANDARD; PRT; 228 AA.
AC Q9KG8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (EC 2.7.7.60)
DE (4-diphosphocytidyl-2C-methyl-D-erythritol synthase) (MEP
cytidyltransferase) (MCT).
GN Name=ispd; OrderedLocusNames=BH0107;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakaseone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- FUNCTION: Catalyzes the formation of 4-diphosphocytidyl-2-C-
methyl-D-erythritol from CTP and 2-C-methyl-D-erythritol 4-
phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: CTP + 2-C-methyl-D-erythritol 4-phosphate =
diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol.
CC -1- PATHWAY: Isoprenoid biosynthesis; isopentenyl-PP biosynthesis via
DXP pathway; isopentenyl-PP from 1-deoxy-D-xylulose 5-phosphate;
step 2.
CC -1- SIMILARITY: Belongs to the ispd family.

This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

EMBL; BA000004; BAB03826.1; -; Genomic DNA.

PIR; C83663; C83663.

HSSP; Q46893; 1I52.

HAMAP; MF_00108; -; 1.

InterPro; IPR008233; DPCME synth.

InterPro; IPR001228; ISPD_synthase.

Pfam; PF01128; ispd; 1.

Pfam; PIRSF006765; DPCME synth; 1.

TIGRFAMs; TIGR00453; ispd; 1.

PROSITE; PS01295; ISPD; 1.

Complete proteome; Isoprene biosynthesis; Nucleotidyltransferase;

Transferase.

SITE 15 15 Involved in transition state

stabilization (By similarity).

SITE 22 22 Involved in transition state

stabilization (By similarity).

SITE 152 152 Positions MEP for the nucleophilic attack

(By similarity).

SITE 208 208 Positions MEP for the nucleophilic attack

(By similarity).

SEQUENCE 228 AA; 25270 MW; 1D94A1361DEDA080 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 228;

Best Local Similarity 100.0%; Pred. No. 3.2e+02;

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 IELGGKP 46
   |||||
Db 25 IELGGKP 31

RESULT 447
ID QBAIG6_BACTN PRELIMINARY; PRT; 231 AA.
AC QBAIG6;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Putative DNA repair protein.
GN OrderedLocusNames=BT3695;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroidetes (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_taxid=818;
RN [1]

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=VPI-5482 / ATCC 29148;
RC MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chang H.C., Hooper L.V., Gordon J.I.;
RL "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
DR EMBL; AE016941; AAO78800.1; -; Genomic_DNA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR001405; RadC.
DR Pfam; PF04002; RadC; 1.
DR ProDom; PD007415; RadC; 1.
DR TIGRFAMs; TIGR00608; radC; 1.
DR PROSITE; PS01302; RADC; UNKNOWN 1.
KW Complete proteome; DNA damage; DNA repair.
SQ SEQUENCE 231 AA; 26254 MW; E251D9E0EAD03B81 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 89 LLQRTAQ 95
   |||||
Db 169 LLQRTAQ 175

RESULT 448
ID Q97MD2_CLOAB PRELIMINARY; PRT; 234 AA.
AC Q97MD2;
DT 01-OCT-2001 (TREMELrel. 18, Created)
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE ABC transporter, ATP-binding protein.
GN OrderedLocusNames=CAC0266;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_taxid=1488;
RN [1]

RN NUCLEOTIDE SEQUENCE.
RP STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RC MEDLINE=21359325; PubMed=11466286;
RX DOI=10.1128/JB.183.16.4823-4838.2001;
RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabatne F., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RJ Bacteriol. 183:4823-4838(2001).
CC -!- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
```

```
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AB007539; AA078247.1; -; Genomic_DNA.
DR PIR; D96932; D96932.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0016887; P:ATPase activity; IEA.
DR GO; GO:0015087; F:cobalt ion transporter activity; IEA.
DR GO; GO:0001678; F:hydrolyase activity; IEA.
DR GO; GO:0000166; P:nucleotide binding; IEA.
DR GO; GO:0006824; P:cobalt ion transport; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transp; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome; Membrane; Nucleotide-binding;
transport.
SQ SEQUENCE 234 AA; 26571 MW; D6624AB7634A7973 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGGLL 90
   |||||
Db 31 GKVLGGLL 37

RESULT 449
ID YOGA_ECOLI STANDARD; PRT; 235 AA.
AC YOGA_ECOLI
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-NOV-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein yqga.
GN Name=yqga; OrderedLocusNames=b2966;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_taxid=562;
RN [1]

RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=KL2 / MG1655;
RC MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
DR EMBL; U28377; AAA69134.1; -; Genomic_DNA.
DR EMBL; U00096; AAC76003.1; -; Genomic_DNA.
DR PIR; B65082; B65082.
DR EcoBASE; EB2811; -.
DR EcoGene; EGI2987; yqga.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 2 22 Potential.
FT TRANSMEM 34 54 Potential.
FT TRANSMEM 56 76 Potential.
FT TRANSMEM 102 122 Potential.
```

```

FT TRANSMEM 147 167 Potential.
FT TRANSMEM 178 198 Potential.
FT TRANSMEM 210 230 Potential.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 5.9%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 12 VLLGGVL 18

RESULT 450
Q83Q81_SHIFL PRELIMINARY; PRT; 235 AA.
AC Q83Q81; Q7C016;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative transport protein.
GN Names=yqgA; OrderedLocusNames=S3166, SF2963;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AB005674; AAN4444.1; -; Genomic_DNA.
DR EMBL; AB016988; AAP18268.1; -; Genomic_DNA.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Complete proteome.
SQ SEQUENCE 235 AA; 24626 MW; 8194D2A93FF9AF31 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 12 VLLGGVL 18

RESULT 451
Q8VS17_KLEOX PRELIMINARY; PRT; 236 AA.
AC Q8VS17;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

```

```

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative repressor protein.
OS Klebsiella oxytoca.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=571;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MSal;
RX PubMed=15257457; DOI=10.1007/s00438-004-1022-8;
RA Shakeri-Garakani A., Brinkkötter A., Schmid K., Turgut S.,
RA Lengeler J.W.;
RT "The genes and enzymes for the catabolism of galactitol, D-tagatose,
RT and related carbohydrates in Klebsiella oxytoca M5al and other enteric
RT bacteria display convergent evolution."
RL Mol. Genet. Genomics 271:717-728(2004).
DR EMBL; AF416702; AAL60164.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001034; HTH_Deor.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00455; Deor; 1.
DR PRINTS; PR00037; HTHLACR.
DR SMART; SM00420; HTH_Deor; 1.
DR PROSITE; PS00894; HTH_Deor_1; 1.
DR PROSITE; PSS1000; HTH_Deor_2; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 236 AA; 25483 MW; 1624D6DBF25F6B8F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 236;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 173 VLLGGVL 179

RESULT 452
Q8YQ77_ANASP PRELIMINARY; PRT; 245 AA.
AC Q8YQ77;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alr3958 protein.
GN OrderedLocusNames=alr3958;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Matanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraiki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; BA000019; BAB75657.1; -; Genomic_DNA.
DR PIR; AG2300; AG2300.
DR InterPro; IPR008538; DUF820.
DR Pfam; PF05685; DUF820; 1.
KW Complete proteome.
SQ SEQUENCE 245 AA; 28544 MW; E7479D921F999138 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 245;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY 45 KPAIVPD 51
Db 67 KPAIVPD 73

RESULT 453
Q63QR8_BURPS
ID Q63QR8_BURPS PRELIMINARY; PRT; 246 AA.
AC Q63QR8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=BFSL2956;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crossett B., Davis P., Deshazer D.,
RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245 (2004).
DR EMBL; BX571965; CAH36966.1; -; Genomic_DNA.
DR InterPro; IPR004382; Cons_Hypoth46.
DR InterPro; IPR006700; DUF558.
DR Pfam; PF04452; DUF558; 1.
DR TIGRFAMs; TIGR00046; Cons_hypoth46; 1.
KW Complete proteome: Hypothetical protein.
SQ SEQUENCE 246 AA; 25867 MW; 8DD14BB4C9507505 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 235 VLAALAA 241

RESULT 454
Q5YXG7_NOCFA
ID Q5YXG7_NOCFA PRELIMINARY; PRT; 247 AA.
AC Q5YXG7;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Putative transcriptional regulator.
GN OrderedLocusNames=nfa22770;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930 (2004).

DR EMBL; AP006618; BAD57124.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000551; HTH_MerR.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00376; MerR; 2.
DR SMART; SM00422; HTH_MER_2.
DR PROSITE; PS0937; HTH_MER_2; 2.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 247 AA; 26633 MW; 2766114CE584C400 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 247;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 202 VLAALAA 208

RESULT 455
Q75CX5_ASHGO
ID Q75CX5_ASHGO PRELIMINARY; PRT; 250 AA.
AC Q75CX5;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE ABR249Wp.
GN Name=ABR249W;
OS Ashbya gossypii (Yeast) (Bremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Bremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegeli S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehmann K., Luedi P., Choi S., Wing R.A., Flavie A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome."
RL Science 304:304-307 (2004)
DR EMBL; AE016815; AAS51022.1; -; Genomic_DNA.
DR AGO; ABR249W; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GDHRDH.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 250 AA; 26726 MW; 7686115867189DF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 219 VLAALAA 225

RESULT 456
Q75DB2_ASHGO
ID Q75DB2_ASHGO PRELIMINARY; PRT; 250 AA.
AC Q75DB2;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)

```

```

DE ABR247MD.
GN Name=ABR247M;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Pohlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippaen P.;
RA "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomycetes cervisiae genome.";
RL Science 304:304-307(2004).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
(CC (SDR) family.
DR EMBL; AE016815; AA851020.1; -; Genomic_DNA.
DR AGD; ABR247M; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR PRINTS; PR00081; GHRDR.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN 1.
KW Complete proteome; Oxidoreductase.
SQ SEQUENCE 250 AA; 26665 MW; PAA56BCA8E179ED5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 250;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
| | | | |
DB 219 VLAALAA 225

RESULT 457
ISPD_RALSO
ID _ISPD_RALSO STANDARD; PRT; 253 AA.
AC QAKYW3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase (EC 2.7.7.60)
DE (4-diphosphocytidyl-2C-methyl-D-erythritol synthase) (MEP
DE cytidyltransferase) (MCT).
GN Name=ispb; OrderedLocNames=RSC1643; ORFNames=RS04018;
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OC NCBI_TaxID=305;
OX [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=GM11000;
RC MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigur P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RT Nature 415:497-502(2002).
RL -1- FUNCTION: Catalyzes the formation of 4-diphosphocytidyl-2-C-
CC methyl-D-erythritol from CTP and 2-C-methyl-D-erythritol 4-
CC phosphate (MEP) (By similarity).
CC -1- CATALYTIC ACTIVITY: CTP + 2-C-methyl-D-erythritol 4-
CC diphosphate + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol.
CC -1- PATHWAY: Isoprenoid biosynthesis; isopentenyl-PP biosynthesis via

```

```

CC DXP pathway; isopentenyl-PP from 1-deoxy-D-xylulose 5-phosphate:
CC step 2.
CC -1- SIMILARITY: Belongs to the ispd family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; AL646065; CAD15345.1; -; Genomic_DNA.
DR HSP; Q46893; IH3M.
DR HAMAP; MF_00108; -.
DR InterPro; IPR008233; DPCME_synth.
DR InterPro; IPR001228; ISPD_synthase.
DR Pfam; PF01128; Ispd; 1.
DR PIRSF; PIRSF006765; DPCME_synth; 1.
DR TIGRPFAMS; TIGR00453; ispd; 1.
DR PROSITE; PS01295; ISPD; 1.
KW Complete proteome; Isoprene biosynthesis; Nucleotidyltransferase;
KW Transferase.
FT SITE 23 23 Involved in transition state
FT SITE 30 30 Involved in transition state
FT SITE 171 171 Stabilization (By similarity).
FT SITE 227 227 Positions MEP for the nucleophilic attack
FT SITE 227 227 Positions MEP for the nucleophilic attack
FT SITE 227 227 Positions MEP for the nucleophilic attack
FT SITE 227 227 Positions MEP for the nucleophilic attack
SQ SEQUENCE 253 AA; 26134 MW; 22CC2D6CFAB3B46E CRC64;

Query Match 5.9%; Score 7; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQRA 93
| | | | |
DB 185 LGLLQRA 191

RESULT 458
QAS101_TETNG
ID QAS101_TETNG PRELIMINARY; PRT; 253 AA.
AC QAS101;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome 5 SCAF14773, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00025788001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygii; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallou O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Cattolico L., Poulain J., De Berardinis V.,
RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McGowan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).

```

```

RN NUCLEOTIDE SEQUENCE.
RP Genoscope; Whitehead Institute Centre for Genome Research;
RG Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
RL -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CHAE01014773; CAG05681.1; -; Genomic_DNA.
FT NON_TER 1
FT NON_TER 253
SQ SEQUENCE 253 AA; 28492 MW; 12F28B3407260DB2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
   |||||
DB 81 LLGGVLA 87

RESULT 459
Q62H02 BURMA PRELIMINARY; PRT; 255 AA.
AC Q62H02;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BMA2472;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Daviden T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarrisa S., Sellengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000010; AAU49676.1; -; Genomic_DNA.
DR TIGR; BMA2472; -.
DR InterPro; IPR004382; Cons_hypoth46.
DR InterPro; IPR006700; DUF558.
DR Pfam; PF04452; DUF558; 1.
DR TIGRFam; TIGR00046; Cons_hypoth46; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 255 AA; 26567 MW; 6C8178403B65205E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 255;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
   |||||
DB 244 VLAALAA 250

RESULT 460
Q81586_9HEPC PRELIMINARY; PRT; 256 AA.
AC Q81586;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE NS4 protein (Fragment).
```

```

OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RA Tsukiyama-Kohara K.;
RT "Antigenicities of group I and II Hepatitis C virus polypeptides-
RT molecular basis of diagnosis.";
RL Virology 0:0-0(1993).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RA Kohara K.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; D14112; BAA03175.1; -; Genomic_RNA.
DR HSSP; P26663; 1CU1.
DR SMR; Q81586; 1-152.
DR InterPro; IPR000745; HCV_NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 1
FT NON_TER 256
SQ SEQUENCE 256 AA; 28069 MW; 6C31B2A9A8D9DE9E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 256;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
   |||||
DB 191 PDKEVLY 197

RESULT 461
HDRE METMA STANDARD; PRT; 259 AA.
ID HDRE METMA STANDARD; PRT; 259 AA.
AC Q8PVM4;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE CoB-CoM heterodisulfide reductase 2 subunit E (EC 1.8.98.1).
GN Name=hdrE; OrderedLocusNames=MM1843;
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wierzer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
RA Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
RN [2]
RN FUNCTION, AND SOURCE OF ELECTRONS.
RC STRAIN=Goel / GoI / ATCC BAA-199 / DSM 3647 / OCM 88;
RX PubMed=9654152; DOI=10.1016/S0014-5793(98)00555-9;
RA Baeumer S., Murakami E., Brodersen J., Gottschalk G., Ragsdale S.W.,
RA Deppenmeier U.;
RT "The F420H2:heterodisulfide oxidoreductase system from Methanosarcina
RT species. 2-Hydroxyphenazine mediates electron transfer from F420H2
RT dehydrogenase to heterodisulfide reductase.";
RL FEBS Lett. 428:295-298(1998).
CC -!- FUNCTION: Part of a complex that catalyzes the reversible
CC reduction of CoM-S-S-CoB to the thiol-coenzymes H-S-CoM (coenzyme
CC M) and H-S-CoB (coenzyme B). HdrE may be responsible for anchoring
CC the complex to the membrane.
CC -!- CATALYTIC ACTIVITY: Coenzyme B + coenzyme M + methanophenazine =
CC N-(7-((2-sulfoethyl)dithio)heptanoyl)-3-O-phospho-L-threonine +
```

```

CC dihydromethanophenazine.
CC -1- PATHWAY: Methanogenesis from acetate; sixth (last) step.
CC -1- PATHWAY: Methanogenesis from carbon dioxide; eighth (last) step.
CC -1- PATHWAY: Methanogenesis from methanol; third (last) step.
CC -1- SUBUNIT: The heterodisulfide reductase 2 is composed of two
CC subunits; hrdR and hrdR (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- MISCELLANEOUS: Methanophenazine seems to mediate electron transfer
CC from a F420-non-reducing hydrogenase to the heterodisulfide
CC reductase 2.
CC -1- SIMILARITY: Belongs to the hrdR family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE013421; AM31539.1; -, Genomic DNA.
CC Complete proteome; Methanogenesis; Oxidoreductase; Transmembrane.
CC TRANSMEM 15 35 Potential.
CC TRANSMEM 104 124 Potential.
CC TRANSMEM 147 167 Potential.
CC TRANSMEM 180 200 Potential.
CC TRANSMEM 218 238 Potential.
CC SEQUENCE 259 AA; 28989 MW; A328ADE9E99406F9 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 259;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 WVLGGV 20
Db 184 WVLGGV 190

RESULT 462
Q4LQ33 9BURK
ID Q4LQ33 9BURK PRELIMINARY; PRT; 259 AA.
AC Q4LQ33;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative resistance protein mnr homolog precursor (Fragment).
GN ORFNames=Scen244DRAFT_3403;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Iserani S., Pitluck S., Richardson P.,
RA "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT H12424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT H12424."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAHL01000030; EAM18093.1; -, Genomic_DNA.
CC Signal.
CC SIGNAL
CC NON_TER <1 35 Potential.

```

```

SQ SEQUENCE 259 AA; 26320 MW; 68D65F5FB066C26 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 26 VLAALAA 32

RESULT 463
Q75DB1 ASHGO
ID Q75DB1 ASHGO PRELIMINARY; PRT; 261 AA.
AC Q75DB1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ASR248WP.
GN Name=ABR248W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Steiner S.,
RA Mehr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome."
RL Science 304:304-307(2004).
CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
CC (SDR) family.
CC EMBL; AE016815; AASS1021.1; -, Genomic_DNA.
CC AGD; ABR248W; -.
CC GO; GO:0016491; F:oxidoreductase activity; IEA.
CC GO; GO:0008152; P:metabolism; IEA.
CC InterPro; IPR002198; ADH_short.
CC InterPro; IPR002347; Adh_short_C2.
CC Pfam; PF00106; adh_short; 1.
CC PRINTS; PR00081; GDRDH.
CC PRINTS; PR00080; SDRFAMILY.
CC PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
CC Complete proteome; Oxidoreductase.
SQ SEQUENCE 261 AA; 27979 MW; FF0338CAEA0CECED CRC64;

Query Match 5.9%; Score 7; DB 2; Length 261;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 230 VLAALAA 236

RESULT 464
THIM STAEF
ID THIM STAEF STANDARD; PRT; 262 AA.
AC Q8CRN9;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Hydroxyethylthiazole kinase (EC 2.7.1.50) (4-methyl-5-beta-
DE hydroxyethylthiazole kinase) (Thz kinase) (TH kinase).
GN Name=thm; OrderedLocNames=SEI691;
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

```

RC STRAIN-ATCC 12228;
 RX DOI=10.1046/j.1365-2958.2003.03671.x;
 RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Pu G., Yang J.,
 RA Qin Z.-Q., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z.,
 RA Yuan Z.-H., Zhao G.-P., Ou D., Danchin A., Wen Y.-M.,
 RT "Genome-based analysis of virulence genes in a non-biofilm-forming
 RT Staphylococcus epidermidis strain (ATCC 12228).";
 RL MOL. MICROBIOL. 49:1577-1593 (2003).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-methyl-5-(2-hydroxyethyl)thiazole =
 CC ADP + 4-methyl-5-(2-phosphoethyl)thiazole.
 CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
 CC -1- PATHWAY: Cofactor biosynthesis; thiamine-PP biosynthesis; 4-
 CC methyl-5-(2-phosphoethyl)thiazole from 4-methyl-5-(2-
 CC hydroxyethyl)thiazole: single step.
 CC -1- SIMILARITY: Belongs to the Thz kinase family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AB016749; AAC05290.1; -; Genomic_DNA.
 DR HSPF; P39593; 1E0Q.
 DR HAWAP; MF_00228; -; 1.
 DR InterPro; IPR000417; Hyethyz_kinase.
 DR Pfam; PF02110; HK; 1.
 DR PIRSF; PIRSF000513; Thz_kinase; 1.
 DR PRINTS; PR01099; HYETHYZKNASE.
 DR TIGRFAMs; TIGR00694; thim; 1.
 KW ATP-binding; Complete proteome; Kinase; Magnesium; Metal-binding;
 KW Nucleotide-binding; Thiamine biosynthesis; Transferase.
 FT METAL 88 88 Magnesium 1 (By similarity).
 FT METAL 120 120 Magnesium 1 (By similarity).
 FT BINDING 39 39 Substrate (via amide nitrogen) (By
 FT similarity).
 FT BINDING 115 115 ATP (beta-phosphate) (By similarity).
 FT BINDING 160 160 ATP (alpha-phosphate) (By similarity).
 FT BINDING 187 187 Substrate (via amide nitrogen) (By
 FT similarity).
 SQ SEQUENCE 262 AA; 28388 MW; A55A929C59B80129 CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 262;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 16 LLGGVLA 22
 |||||
 Db 191 LLGGVLA 197
 RESULT 465
 ID THIM STAEQ STANDARD; PRT; 262 AA.
 AC QSHMC9;
 DT 13-SEP-2005 (Rel. 48, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DE Hydroxyethylthiazole kinase (EC 2.7.1.50) (4-methyl-5-beta-
 DE hydroxyethylthiazole kinase) (Thz kinase) (TH kinase).
 DN Name=thim; OrderedLocNames=SERP1699;
 OS Staphylococcus epidermidis (strain ATCC 35984 / RP62A).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=176279;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=HTA426;
 RC PubMed=15774886; DOI=10.1128/JB.187.7.2426-2438.2005;
 RA Gill S.R., Fouts D.E., Archer G.L., Mongodin E.F., DeBoy R.T.,
 RA Ravel J., Paulsen I.T., Kolony J.F., Brinkac L.M., Beanan M.J.,
 RA Dodson R.J., Daugherty S.C., Madupu R., Angluoli S.V., Durkin A.S.,
 RA Haft D.H., Vamathevan J.J., Khouri H., Utterback T.R., Lee C.,

RA Dimitrov G., Jiang L., Qin H., Weidman J., Tran K., Kang K.H.,
 RA Hance I.R., Nelson K.E., Fraser C.M.;
 RT "Insights on evolution of virulence and resistance from the complete
 RT genome analysis of an early methicillin-resistant Staphylococcus
 RT aureus strain and a biofilm-producing methicillin-resistant
 RT Staphylococcus epidermidis strain.";
 RL J. Bacteriol. 187:2426-2438 (2005).
 CC -1- CATALYTIC ACTIVITY: ATP + 4-methyl-5-(2-hydroxyethyl)thiazole =
 CC ADP + 4-methyl-5-(2-phosphoethyl)thiazole.
 CC -1- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
 CC -1- PATHWAY: Cofactor biosynthesis; thiamine-PP biosynthesis; 4-
 CC methyl-5-(2-phosphoethyl)thiazole from 4-methyl-5-(2-
 CC hydroxyethyl)thiazole: single step.
 CC -1- SIMILARITY: Belongs to the Thz kinase family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; CP0000029; AAW55077.1; -; Genomic_DNA.
 DR TIGR; SERP1699; -; 1.
 DR HAWAP; MF_00228; -; 1.
 DR InterPro; IPR000417; Hyethyz_kinase.
 DR Pfam; PF02110; HK; 1.
 DR PIRSF; PIRSF000513; Thz_kinase; 1.
 DR PRINTS; PR01099; HYETHYZKNASE.
 DR TIGRFAMs; TIGR00694; thim; 1.
 KW ATP-binding; Complete proteome; Kinase; Magnesium; Metal-binding;
 KW Nucleotide-binding; Thiamine biosynthesis; Transferase.
 FT METAL 88 88 Magnesium 1 (By similarity).
 FT METAL 120 120 Magnesium 1 (By similarity).
 FT BINDING 39 39 Substrate (via amide nitrogen) (By
 FT similarity).
 FT BINDING 115 115 ATP (beta-phosphate) (By similarity).
 FT BINDING 160 160 ATP (alpha-phosphate) (By similarity).
 FT BINDING 187 187 Substrate (via amide nitrogen) (By
 FT similarity).
 SQ SEQUENCE 262 AA; 28358 MW; A55A928918EC0129 CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 262;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 16 LLGGVLA 22
 |||||
 Db 191 LLGGVLA 197
 RESULT 466
 ID Q5LOJ7 GEOKA PRELIMINARY; PRT; 264 AA.
 AC Q5LOJ7;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE CDP-diglyceride synthase(Phosphatidate cytidyltransferase)
 DE (EC 2.7.7.41).
 DN OrderedLocNames=GK1254;
 OS Geobacillus kaustophilus.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
 OX NCBI_TaxID=1462;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=HTA426;
 RC PubMed=15576355; DOI=10.1093/nar/gkh970;
 RA Takami H., Takaki Y., Chee G.-J., Nishi S., Shimamura S., Suzuki H.,
 RA Matsui S., Uchiyama I.;
 RT "Thermoadaptation trait revealed by the genome sequence of
 RT thermophilic Geobacillus kaustophilus.";
 RL Nucleic Acids Res. 32:6292-6303 (2004).

DR EMBL; BA000043; BAD75539.1; -; Genomic DNA.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0004605; F: phosphatidate cytidyltransferase activity; IEA.
 DR GO; GO:0016740; F: transferase activity; IEA.
 DR GO; GO:0008654; P: phospholipid biosynthesis; IEA.
 DR InterPro; IPR000374; PC trans.
 DR PANTHER; PTHR13773; PC trans.
 DR Pfam; PF01148; CTP transf_1; 1.
 DR PROSITE; PS01315; CDS; 1.
 DR Complete proteome; Nucleotidyltransferase; Transferrase; Transmembrane.
 SQ SEQUENCE 264 AA; 28156 MW; 11C7BD3C4D6CECC4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 264;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
 |||||
 Db 83 GVLAALA 89

RESULT 467
 Q4NSY4_9DEL
 ID Q4NSY4_9DEL PRELIMINARY; PRT; 267 AA.
 AC Q4NSY4;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Similar to Membrane-anchored protein predicted to be involved in
 DE regulation of amyloplullulanase precursor.
 GN ORFNames=AdenDRAFT_1687;
 OS Anaeromyxobacter dehalogenans 2CP-C.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophactereinae; Myxococcaceae; Anaeromyxobacter.
 OX NCBI_TaxID=290397;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RA "Sequencing of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RA "Annotation of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAHD01000020; EAL78648.1; -; Genomic DNA.
 SQ SEQUENCE 267 AA; 28598 MW; 77B1E167943156DA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 6 VLAALAA 12

RESULT 468
 Q4LUAS_9BURK
 ID Q4LUAS_9BURK PRELIMINARY; PRT; 272 AA.

Q4LUAS;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE AzlC protein.
 GN ORFNames=Bcen2424DRAFT_4492;
 OS Burkholderia cenocepacia HI2424.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
 OX NCBI_TaxID=331272;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HI2424;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RA "Sequencing of the draft genome assembly of Burkholderia cenocepacia
 RT HI2424."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=HI2424;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RA "Annotation of the draft genome assembly of Burkholderia cenocepacia
 RT HI2424."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAHL01000015; EAM19572.1; -; Genomic DNA.
 SQ SEQUENCE 272 AA; 29519 MW; 85666A1085BIA243 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 272;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 239 VLAALAA 245

RESULT 469
 Q4NTJ8_9DEL
 ID Q4NTJ8_9DEL PRELIMINARY; PRT; 274 AA.
 AC Q4NTJ8;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein precursor.
 GN ORFNames=AdenDRAFT_2044;
 OS Anaeromyxobacter dehalogenans 2CP-C.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophactereinae; Myxococcaceae; Anaeromyxobacter.
 OX NCBI_TaxID=290397;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RA "Sequencing of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=2CP-C;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RA "Annotation of the draft genome assembly of Anaeromyxobacter
 RT dehalogenans 2CP-C."; to the EMBL/GenBank/DBJ databases.
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an

```

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA001000017; EAL78929.1; -; Genomic_DNA.
KW Hypothetical protein; Signal.
FT SIGNAL 1 22 Potential.
SQ SEQUENCE 274 AA; 28856 MW; E3B4CE2A86099599 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 274;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 VLAALAA 26
   |||||
Db 10 VLAALAA 16

RESULT 470
ID Q7P114 CHRV0 PRELIMINARY; PRT; 275 AA.
AC Q7P114;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
DE OrderedLocusNames=CV0400;
GN Chromobacterium violaceum.
OS Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.P., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Araujo J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataua L.A.M.,
RA Batista J.S., Belo A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brigidio M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Carago A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascardo J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Pantinatti F., Farias I.P., Felipe M.S.A., Furlan L.R.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furlan L.R.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grissard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.P., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.Q., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meisner R.V., Moreira M.A.M.,
RA Nascimento F.F., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.P.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos P.R., Schneider M.P.C., Seuneez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vettore A., Wassem R., Zaba A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
DR EMBL; AE016911; AAQ58078.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. ...); IEA.
DR GO; GO:0005381; F:iron ion transporter activity; IEA.
DR GO; GO:0008565; F:protein transporter activity; IEA.
DR GO; GO:0006826; P:iron ion transport; IEA.
DR GO; GO:0015031; P:protein ion transport; IEA.
DR InterPro; IPR001719; AP_endonuclease2.
DR InterPro; IPR003538; TonB.
DR Pfam; PF03544; TonB; 1.
DR PRINTS; PR01374; TONBPROTEIN.

DR TIGRFAMs; TIGR01352; tonB_Cterm; 1.
DR PROSITE; PS00729; AP_NUCLEASE_F2.1; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 275 AA; 28690 MW; 5DB7706AF496B951 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 LGGVLA 23
   |||||
Db 56 LGGVLA 62

RESULT 471
ID SNAB STRPR STANDARD; PRT; 276 AA.
AC P54933;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Pristinamycin IIA synthase subunit B (PIIA synthase subunit B).
GN Name=snaB;
OS Streptomyces pristinaespiralis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=38300;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SF92;
RX MEDLINE=95394837; PubMed=7665509;
RA Blanc V., Lagneau D., Didier P., Gil P., Lacroix P., Crouzet J.;
RT "Cloning and analysis of structural genes from Streptomyces
RT pristinaespiralis encoding enzymes involved in the conversion of
RT pristinamycin IIB to pristinamycin IIA (PIIA): PIIA synthase and
RT NADH:riboflavin 5'-phosphate oxidoreductase.";
RL J. Bacteriol. 177:5206-5214(1995).
RN [2]
RP PROTEIN SEQUENCE OF 1-22 AND 121-135.
RX MEDLINE=95394836; PubMed=7665508;
RA Thibaut D., Ratet N., Bisch D., Faucher D., Debussche L., Blanche P.;
RT "Purification of the two-enzyme system catalyzing the oxidation of the
RT D-proline residue of pristinamycin IIB during the last step of
RT pristinamycin IIA biosynthesis.";
RL J. Bacteriol. 177:5199-5205(1995).
CC -1- FUNCTION: Catalyzes the oxidation of the proline residue of
CC pristinamycin IIB (PIIB) to pristinamycin IIA (PIIA).
CC -1- COFACTOR: FMN.
CC -1- SUBUNIT: Heterodimer of two subunits, snaA and snaB.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U21215; AAA83565.1; -; Genomic_DNA.
DR InterPro; IPR011253; Luciferase_like.
KW Direct protein sequencing; Flavoprotein; FMN; Monooxygenase;
KW Oxidoreductase.
FT INIT MET 0
SQ SEQUENCE 276 AA; 28633 MW; EA8EA9E738D02E45 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 VLAALAA 26
   |||||
Db 239 VLAALAA 245

RESULT 472

```

Q8ZWH5_PYRAE
ID Q8ZWH5_PYRAE PRELIMINARY; PRT; 276 AA.
AC Q8ZWH5_2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein PAE1784.
GN OrderedLocusNames=PAE1784;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=IN2 / ATCC 51768 / DSM 7523;
RX MEDLINE=2164397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum".
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AE009839; AAL63727.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000537; UbiA_prenyltrans.
DR Pfam; PF01040; UbiA; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 276 AA; 2870 MW; 24E2F57F671HDPF8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 276;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 12 GVLAALA 18
|||||

RESULT 473
Q83SW1_SALTI
ID Q83SW1_SALTI PRELIMINARY; PRT; 277 AA.
AC Q83SW1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=t3172;
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RX DOI=10.1128/JB.185.7.2330-2337.2003;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT and CT18".
RL J. Bacteriol. 185:2330-2337(2003).
DR EMBL; AB016844; AA070711.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001034; HTH_Deor.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR Pfam; PF00455; Deor; 1.
DR PRINTS; PR00037; HTHLACR.
DR SMART; SM00420; HTH_Deor; 1.
DR PROSITE; PS00894; HTH_Deor 1; 1.
DR PROSITE; PS1000; HTH_Deor 2; 1.
KW DNA-binding; Hypothetical protein; Transcription;
Transcription regulation.

SQ SEQUENCE 277 AA; 30189 MW; 3797CE43AD9CCD4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 147 VLLGGVL 153
|||||

RESULT 474
Q8ZLV5_SALTY
ID Q8ZLV5_SALTY PRELIMINARY; PRT; 277 AA.
AC Q8ZLV5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Aga operon transcriptional repressor.
GN Name=agar; OrderedLocusNames=STM3252;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L72;
RX MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT L72".
RL Nature 413:852-856(2001).
DR EMBL; AB008850; AAL22124.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR011991; Wing_hlx_DNA_bd.
DR InterPro; IPR011991; HTH_Deor.
DR Pfam; PF00455; Deor; 1.
DR PRINTS; PR00037; HTHLACR.
DR SMART; SM00420; HTH_Deor; 1.
DR PROSITE; PS00894; HTH_Deor 1; 1.
DR PROSITE; PS1000; HTH_Deor 2; 1.
KW Complete proteome; DNA-binding; Transcription;
Transcription regulation.
SQ SEQUENCE 277 AA; 30156 MW; 7268DEA3B79CDSDF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 147 VLLGGVL 153
|||||

RESULT 475
Q5PL87_SALPA
ID Q5PL87_SALPA PRELIMINARY; PRT; 277 AA.
AC Q5PL87;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein agar.
GN Name=agar; OrderedLocusNames=SPA3121;
OS Salmonella paratyphi-a.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=54388;

DR	PROSITE; PS00194; THIOREDOXIN; 1.
SQ	SEQUENCE 278 AA; 30460 MW; DB559EC5926F5B7E CRC64;
Query Match	5.9%; Score 7; DB 2; Length 278;
Best Local Similarity	100.0%; Pred.No. 3.7e+02;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	19 GVLAALA 25
Db	87 GVLAALA 93
RESULT 477	
ID	Q8DAI9_VIBVU PRELIMINARY; PRT; 279 AA.
AC	Q8DAI9;
DT	01-MAR-2003 (T-EMBLrel. 23, Created)
DT	01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE	AraC-type DNA-binding domain-containing protein.
GN	OrderedLocusNames=VV12208;
OS	Vibrio vulnificus.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Vibrio.
OX	NCBI_TaxID=672;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RC	STRAIN=CMCP6;
RA	Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA	Choy H.E.;
RL	"Complete genome sequence of Vibrio vulnificus CMCP6.";
RT	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
SR	EMBL; AE016804; AA010590.1; -; Genomic_DNA.
DR	GO; GO:0005622; C:intracellular; IEA.
DR	GO; GO:0003700; P:transcription factor activity; IEA.
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro; IPR003313; Arac bd.
DR	InterPro; IPR012287; Homeodomain-rel.
DR	InterPro; IPR000005; HTHArac.
DR	Pfam; PF00165; HTH_Arac; 2.
DR	PRINTS; PR00032; HTHARAC.
DR	SMART; SMO0342; HTH_ARAC; 1.
DR	PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_2.
DR	PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW	Activator; Complete proteome; DNA-binding; Transcription;
KW	Transcription regulation.
SQ	SEQUENCE 279 AA; 30433 MW; E5181E24D12CEA37 CRC64;
Query Match	5.9%; Score 7; DB 2; Length 279;
Best Local Similarity	100.0%; Pred.No. 3.7e+02;
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	19 GVLAALA 25
Db	160 GVLAALA 166
RESULT 478	
ID	Q7MJL2_VIBVY PRELIMINARY; PRT; 279 AA.
AC	Q7MJL2;
DT	01-MAR-2004 (T-EMBLrel. 26, Created)
DT	01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DT	01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE	AraC-type DNA-binding domain-containing protein.
GN	OrderedLocusNames=VV2149;
OS	Vibrio vulnificus (strain YJ016).
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC	Vibrionaceae; Vibrio.
OX	NCBI_TaxID=196600;
RN	[1]
RP	NUCLEOTIDE SEQUENCE.
RX	PubMed; 14656965; DOI=10.1101/gr.1295503;

RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-P.,
RT "Comparative genome analysis of *Vibrio vulnificus*, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC94913.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003313; Arac bd.
DR InterPro; IPR012287; Homeodomain-rel.
DR InterPro; IPR000005; HTHArac.
DR Pfam; PF00165; HTH_Arac; 2.
DR PRINTS; PR00032; HTHARAC.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_2.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW Activator; Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 279 AA; 30433 MW; E5181E24D12CEA37 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
|||||
DB 160 GVLAALA 166

RESULT 479
Q83YD5_STRHY
ID Q83YD5_STRHY PRELIMINARY; PRT; 280 AA.
AC Q83YD5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Shy15.
GN Names:shy15;
OS Streptomyces hygroscopicus subsp. yingchengensis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=228732;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Qiu L., Wang H., Wu Z., Lu Y.T., Deng Z., Zhao G.;
RT "Identification and function analysis of a MDR efflux protein-Hmr19,
RT in a sequenced genomic DNA fragment from Streptomyces hygroscopicus
RT yingchengensis 10-22.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AY260760; AAP2162.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD_transp.
DR Pfam; PF00528; BPD_transp_1; 1.
DR PROSITE; PS05028; ABC_TM1; 1.
KW Transmembrane; Transport.
SQ SEQUENCE 280 AA; 28990 MW; 7BE3D5865661304C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||

DB 169 VLAALAA 175

RESULT 480
Q93J38_STRCO
ID Q93J38_STRCO PRELIMINARY; PRT; 284 AA.
AC Q93J38;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative integral membrane protein.
GN OrderedLocNames=SCQ3968; ORFNames=SCBAC25E3.05c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Batsman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphoy L.D., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL393118; CAC4692.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 284 AA; 28215 MW; 3025B2BD85958D91 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
|||||
DB 186 VLAALAA 192

RESULT 481
Q6A624_PROAC
ID Q6A624_PROAC PRELIMINARY; PRT; 284 AA.
AC Q6A624;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ABC transporter, putative permease.
GN OrderedLocNames=PPA2081;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.
OX NCBI_TaxID=1747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Bruggemann H., Henne A., Hoster F., Liesegang H., Wietzer A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of *Propionibacterium acnes*, a commensal
RT of human skin.";
RL Science 305:671-673(2004).
DR EMBL; AE017283; AAT83789.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0042266; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011527; ABC_membrane_1.

DR PROSITE; PS50929; ABC_TM1F; 1.
 KW Complete proteome.
 SQ SEQUENCE 284 AA; 30458 MW; DC75AB6C4CD14E5F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 284;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 77 VLAALAA 83

RESULT 482

Q8G2Y2 BRUSU
 ID Q8G2Y2 BRUSU PRELIMINARY; PRT; 285 AA.
 AC Q8G2Y2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Tetrapyrrole methylase family protein.
 GN OrderedLocusNames=BR0177;
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 Read T.D., Dodson R.J., Unayam L.A., Brinkac L.M., Beanan M.J.,
 Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
 Van Aken S.E., Biedmuller S., Tettelin H., Gill S.R., White O.,
 Salzberg S.L., Hoover D.L., Lindler L.E., Helling S.M., Boyle S.M.,
 Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 animal and plant pathogens and symbionts."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
 DR EMBL; AE014291; AAN29130.1; -; Genomic_DNA.
 DR TIGR; BR0177; -.
 DR GO; GO:0008168; P:methyltransferase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000878; Cor/por Mettransf.
 DR InterPro; IPR008189; UPF0011-
 DR Pfam; PF00590; TP_methylase; 1.
 DR PIRSF; PIRSF005917; Mcase_Yral; 1.
 DR TIGRFAMs; TIGR000096; UPF0011; 1.
 DR PROSITE; PS01296; UPF0011; 1.
 KW Complete proteome; Methyltransferase.
 SQ SEQUENCE 285 AA; 30064 MW; FIA8P71687953EE9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 285;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 121 VLAALAA 127

RESULT 483

Q8PVM8 METWA
 ID Q8PVM8 METWA PRELIMINARY; PRT; 286 AA.
 AC Q8PVM8;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Acetyltransferase (EC 2.3.1.-).
 GN OrderedLocusNames=MM1935;
 OS Methanosarcina mazei (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=2212082; PubMed=12125824;
 RA Depemmer U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
 Martinez-Arias R., Henne A., Wierzer A., Baeumer S., Jacobi C.,
 Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,
 Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
 Fritz H.-J., Gottschalk G.;
 RT "The genome of Methanosarcina mazei: evidence for lateral gene
 transfer between Bacteria and Archaea."
 RL J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).
 DR EMBL; AE013430; AAM31631.1; -; Genomic DNA.
 DR GO; GO:0008415; P:acyltransferase activity; IEA.
 DR GO; GO:0008080; P:N-acetyltransferase activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR InterPro; IPR000182; GCN5acetyl_trans.
 DR Pfam; PF00583; Acetyltransf_1.
 KW Acyltransferase; Complete proteome; Transferase.
 SQ SEQUENCE 286 AA; 32676 MW; 976B5F7D41F841AF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 286;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47
 |||||
 Db 55 ELGGKPA 61

RESULT 484

Q9RKH3 STRCO
 ID Q9RKH3 STRCO PRELIMINARY; PRT; 287 AA.
 AC Q9RKH3;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein SCF46.16C.
 GN OrderedLocusNames=SCO3459; ORFNames=SCB46.16C;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cardeno-Tarraga A.-M., Challis G.L.,
 Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147 (2002).
 DR EMBL; AL039116; CAB61867.1; -; Genomic DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 287 AA; 30771 MW; F48CA3B886E17AEF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 287;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
 |||||
 Db 159 GGVLAAL 165

```

RESULT 485
Q6BMQ0 DEBHA
ID Q6BMQ0_DEBHA PRELIMINARY; PRT; 288 AA.
AC Q6BMQ0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to CAGL0J11198g Candida glabrata.
GN OrderedLocNames=DEHA0F03938g;
OS Debaryomyces hansenii (Yeast) (Torulaspora hansenii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Debaryomycetes.
OX NCBI_TaxID=4959;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 36219 / CBS 767;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barney S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,
RA Boistrame A., Boyer J., Catolico L., Confanioli F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Niclaud J.-M., Nikolski M., Oztas S., Oxier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
CC Nature 430:35-44 (2004).
CC -1- SIMILARITY: Belongs to the ribosomal protein L22P family.
DR EMBL; CR382138; CAG88834.1; -; Genomic_DNA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR001063; Ribosomal_L22.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
KW Complete proteome; Ribonucleoprotein; Ribosomal protein.
SQ SEQUENCE 288 AA; 33787 MW; C71E50332C7CB17D CRC64;

Query Match 5.9%; Score 7; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQRA 93
DB 5 LGLLQRA 11
|||||

RESULT 486
Q6ACJ2 LEIXX
ID Q6ACJ2_LEIXX PRELIMINARY; PRT; 288 AA.
AC Q6ACJ2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ABC transporter, permease protein.
GN Name=rbcC; OrderedLocNames=Lxx22250;
OS Leifsonia xylii (subsp. xylii).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Microbacteriaceae; Leifsonia.
OX NCBI_TaxID=59736;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CTCB07;
RX PubMed=15305603;
RA Monteiro-Vitorello C.B., Camargo L.E.A., Van Sluys M.A.,
RA Kitajima J.P., Truffi D., do Amaral A.M., Harakava R.,
de Oliveira J.C.F., Wood D., de Oliveira M.C., Miyaki C.Y.,
Takita M.A., da Silva A.C.R., Furlan L.R., Carraro D.M., Camarotte G.,
Almeida N.F. Jr., Carrier H., Coutinho L.L., El-Dorri H.A.,
Ferro M.I.T., Gagliardi P.R., Gigliotti E., Goldman M.H.S.,
Goldman G.H., Kimura E.T., Ferro S.M.Z., Kuramae E.E., Lemos E.G.M.,
Lemos M.V.F., Mauro S.M.Z., Machado S.A., Marino C.L., Menck C.F.,
Nunes L.R., Oliveira R.C., Pereira G.G., Siqueira W., de Souza A.A.,
Tsai S.M., Zanca A.S., Simpson A.J.G., Brumbley S.M., Setubal J.C.;
RT "The genome sequence of the Gram-positive sugarcane pathogen Leifsonia
xylii subsp. xylii.";
RL Mol. Plant Microbe Interact. 17:827-836 (2004).
RE EMBL; AE016822; AAT89901.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 1.
KW Complete proteome.
SQ SEQUENCE 288 AA; 28067 MW; 56CB560AC07FIDPEC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 288;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 23
DB 48 LGGVLA 54
|||||

RESULT 487
Q5LLG7 SILPO
ID Q5LLG7_SILPO PRELIMINARY; PRT; 288 AA.
AC Q5LLG7;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Transcriptional regulator, AraC family.
GN OrderedLocNames=SQ0A0060;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C., Eisen J.A.,
RA Brinkac L.M., Lewis M., Johnson S., Weaver B., Pai G., Eisen J.A.,
RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
marine environment.";
RL Nature 432:910-913 (2004).
RE EMBL; CP000032; AA97200.1; -; Genomic_DNA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR000005; HTHARAC.
DR InterPro; IPR002016; Peroxidase.
DR Pfam; PF00165; HTH_ARAC; 2.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN 1.
KW Activator; Complete proteome; DNA-binding; Plasmid; Transcription;
Transcription regulation.
SQ SEQUENCE 288 AA; 31361 MW; 8CD5D9EF370BAP28 CRC64;

```

Query Match 5.9%; Score 7; DB 2; Length 288;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 LGLLQRA 93
 |||||
 DB 170 LGLLQRA 176

RESULT 488
 O67RA6 SYMTH
 ID O67RA6 SYMTH PRELIMINARY; PRT; 289 AA.
 AC O67RA6;
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE Branched-chain amino acid ABC transporter permease protein.
 DB OrderedLocusNames=STH802;
 GN Symbiobacterium thermophilum.
 OS Symbiobacterium thermophilum.
 OC Bacteria; Actinobacteria; Symbiobacterium.
 OX NCBI_TaxID=2734;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IAM14863;
 RX PubMed=15383646; DOI=10.1093/nar/gkh830;
 RA Ueda K., Yamashita A., Ishikawa J., Shinada M., Watsui T.,
 RA Morimura K., Ikeda H., Hattori M., Beppu T.;
 RT "Genome sequence of Symbiobacterium thermophilum, an uncultivable
 bacterium that depends on microbial commensalism.";
 RL Nucleic Acids Res. 32:4937-4944(2004).
 DR EMBL; AP066840; BAD39787.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001851; P:transport; IEA.
 DR Pfam; PF02653; BPD_transp_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 289 AA; 30513 MW; C2CF632C3D42F494 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 289;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 LGGVLA 23
 |||||
 DB 194 LGGVLA 200

RESULT 489
 Q89G69 BRAJA
 ID Q89G69 BRAJA PRELIMINARY; PRT; 293 AA.
 AC Q89G69;
 DT 01-JUN-2003 (TRENBLrel. 24, Created)
 DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
 DE ABC transporter permease protein.
 GN OrderedLocusNames=bll16476;
 OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 OX NCBI_TaxID=375;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
 RA Sasamoto S., Watanabe A., Idegawa K., Iriiguchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimpo S., Teurloka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 Bradyrhizobium japonicum USDA110.";
 RL DNA Res. 9:189-197(2002).
 DR EMBL; BA000040; BAC51741.1; -; Genomic_DNA.

DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001851; Bac_inmem_transp.
 DR Pfam; PF02653; BPD_transp_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 293 AA; 31718 MW; E6D86D06B237F0B4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 293;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 LGGVLA 22
 |||||
 DB 21 LGGVLA 27

RESULT 490
 AMPM_PYRFU
 ID AMPM_PYRFU STANDARD; PRT; 295 AA.
 AC P56218;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 13-SEP-2005 (Rel. 46, Last annotation update)
 DE Methionine aminopeptidase (EC 3.4.11.18) (MAP) (Peptidase M).
 GN Name=nacp; OrderedLocusNames=PF0541;
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP PROTEIN SEQUENCE OF 1-35, CHARACTERIZATION, AND MASS SPECTROMETRY.
 RC STRAIN=VCI / DSM 3638 / ATCC 43587 / JCM 8422;
 RX PubMed=9399590;
 RA Tsunasawa S., Izu Y., Miyagi M., Kato I.;
 RT "Methionine aminopeptidase from the hyperthermophilic Archaeon
 Pyrococcus furiosus: molecular cloning and overexpression in
 Escherichia coli of the gene, and characteristics of the enzyme.";
 RL J. Biochem. 122:843-850(1997).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RX MEDLINE=98234705; PubMed=9573622; DOI=10.1006/jmbi.1997.3940;
 RA Tahirov T.H., Oki H., Tsukihara T., Ogasahara K., Yutani K.,
 RA Libeu C.P., Izu Y., Tsunasawa S., Kato I.;
 RT "High-resolution crystals of methionine aminopeptidase from Pyrococcus
 furiosus obtained by water-mediated transformation.";
 RL J. Struct. Biol. 121:68-72(1998).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RX MEDLINE=99030464; PubMed=9811545; DOI=10.1006/jmbi.1998.2146;
 RA Tahirov T.H., Oki H., Tsukihara T., Ogasahara K., Yutani K.,
 RA Izu Y., Tsunasawa S., Kato I.;
 RT "Crystal structure of methionine aminopeptidase from hyperthermophile,
 Pyrococcus furiosus.";
 RL J. Mol. Biol. 284:101-124(1998).
 CC -1- FUNCTION: Removes the amino-terminal methionine from nascent
 proteins.
 CC -1- CATALYTIC ACTIVITY: Release of N-terminal amino acids,
 preferentially methionine, from peptides and arylamides.
 CC -1- COFACTOR: Binds 2 cobalt ions per subunit.
 CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
 pH dependence:
 Optimum pH is 7-8;
 Temperature dependence:
 Optimum temperature is about 90 degrees Celsius;
 -1- SUBUNIT: Monomer.
 CC -1- MASS SPECTROMETRY: MW=32848; METHOD=Electrospray; RANGE=1-295;

```
CC CC NOTE=Ref.2.
CC CC -1- SIMILARITY: Belongs to the peptidase M24C family.
CC CC -----
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use as long as its content is in no way modified and this statement is not
CC CC removed.
CC CC -----
CC CC EMBL; AE010177; AAL80665.1; -; Genomic_DNA.
CC CC PIR; JC5671; JC5671.
CC CC PDB; 1WKM; X-ray; A/B=1-295.
CC CC PDB; 1XGM; X-ray; A/B=1-295.
CC CC PDB; 1XGN; X-ray; A/B=1-295.
CC CC PDB; 1XGO; X-ray; A/B=1-295.
CC CC PDB; 1XGS; X-ray; A/B=1-295.
CC CC MEROPS; M24.002; -.
CC CC InterPro; IPR001714; Pept_M24_MAP.
CC CC InterPro; IPR002468; Pept_M24A_MAP2.
CC CC InterPro; IPR000994; Peptidase_M24.
CC CC PANTHER; PTHR10804; Peptidase_M24; 1.
CC CC Pfam; PF00557; MAPEPTIDASE.
CC CC PRINTS; PR00599; MAPEPTIDASE.
CC CC TIGRFAMs; TIGR00501; met_pdase_II; 1.
CC CC PROSITE; PS01202; MAP_2; 1.
CC CC 3D-structure; Aminopeptidase; Cobalt; Complete proteome;
CC CC Direct protein sequencing; Hydrolase; Metal-binding; Protease.
CC CC METAL 82 82
CC CC METAL 93 93
CC CC METAL 153 153
CC CC METAL 187 187
CC CC METAL 280 280
CC CC HELIX 3 23
CC CC TURN 26 27
CC CC STRAND 29 29
CC CC HELIX 30 43
CC CC TURN 44 45
CC CC STRAND 47 48
CC CC STRAND 53 56
CC CC TURN 57 58
CC CC STRAND 59 61
CC CC TURN 67 68
CC CC STRAND 72 72
CC CC TURN 75 76
CC CC STRAND 78 87
CC CC TURN 88 89
CC CC STRAND 90 99
CC CC TURN 100 101
CC CC HELIX 106 121
CC CC TURN 124 125
CC CC STRAND 127 127
CC CC TURN 128 128
CC CC HELIX 129 140
CC CC TURN 141 143
CC CC STRAND 145 146
CC CC TURN 148 149
CC CC STRAND 152 154
CC CC TURN 156 156
CC CC TURN 157 158
CC CC STRAND 159 159
CC CC STRAND 165 166
CC CC TURN 172 173
CC CC TURN 177 177
CC CC TURN 180 181
CC CC STRAND 183 186
CC CC STRAND 189 191
CC CC STRAND 197 208
CC CC TURN 217 229
CC CC TURN 230 232
CC CC STRAND 235 236
CC CC HELIX 238 240
CC CC TURN 241 243
CC CC HELIX 246 259
```

```
FT TURN 260 260
FT STRAND 262 270
FT TURN 271 272
FT STRAND 276 278
FT STRAND 280 285
FT STRAND 290 292
FT TURN 293 295
SQ SEQUENCE 295 AA; 32842 MW; 9739BC55F812B65B CRC64;

Query Match 5.9%; Score 7; DB 1; Length 295;
Best Local Similarity 100.0%; Pred.No.3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47
Db 43 ELGGKPA 49

RESULT 491
AMPM_PYRHO STANDARD; PRT; 295 AA.
AC OS8362;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Methionine aminopeptidase (EC 3.4.11.18) (MAP) (Peptidase M).
GN Name=map; OrderedLocNames=PH0628;
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Koushi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb P.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT *Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.*;
RL DNA Res. 5:55-76(1998).
CC -1- FUNCTION: Removes the amino-terminal methionine from nascent
CC proteins (By similarity).
CC -1- CATALYTIC ACTIVITY: Release of N-terminal amino acids,
CC preferentially methionine, from peptides and arylamides.
CC -1- COFACTOR: Binds 2 cobalt ions per subunit (By similarity).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SIMILARITY: Belongs to the peptidase M24C family.
CC -----
CC CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use as long as its content is in no way modified and this statement is not
CC CC removed.
CC CC -----
CC CC EMBL; BA000001; BAA29717.1; -; Genomic_DNA.
CC CC PIR; C71107; C71107.
CC CC HGSP; P56218; 1XGS.
CC CC SMR; O58362; 1-295.
CC CC MEROPS; M24.002; -.
CC CC InterPro; IPR001714; Pept_M24_MAP.
CC CC InterPro; IPR002468; Pept_M24A_MAP2.
CC CC InterPro; IPR000994; Peptidase_M24.
CC CC PANTHER; PTHR10804; Peptidase_M24; 1.
CC CC Pfam; PF00557; Peptidase_M24; 1.
CC CC PRINTS; PR00599; MAPEPTIDASE.
CC CC TIGRFAMs; TIGR00501; met_pdase_II; 1.
CC CC PROSITE; PS01202; MAP_2; 1.
CC CC Aminopeptidase; Cobalt; Complete proteome; Hydrolase; Metal-binding;
CC CC Protease.
KW
```

```
FT METAL 82 82 Cobalt 2 (By similarity).
FT METAL 93 93 Cobalt 1 and 2 (By similarity).
FT METAL 153 153 Cobalt 1 (By similarity).
FT METAL 187 187 Cobalt 1 (By similarity).
FT METAL 280 280 Cobalt 1 and 2 (By similarity).
SQ SEQUENCE 295 AA; 32795 MW; D228FA377CEB2AAC CRC64;

Query Match 5.9%; Score 7; DB 1; Length 295;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 ELGGKPA 47
| | | | |
Db 43 ELGGKPA 49

RESULT 492
Q5JGD1_PYRKO PRELIMINARY; PRT; 295 AA.
AC Q5JGD1;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Methionyl aminopeptidase.
GN OrderedLocusNames=TKL183;
OS Pyrococcus kodakaraensis (Thermococcus kodakaraensis).
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Thermococcus.
OC NCBI_TaxID=69014;
RN [1];
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KOD1;
RX PubMed=15710748; DOI=10.1101/gr.3003105;
RA Fukui T., Atomi H., Kanai T., Mateumi R., Fujiwara S., Imanaka T.;
RT "Complete genome sequence of the hyperthermophilic archaeon
RT Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus
RT genomes.";
RL Genome Res. 15:352-363(2005).
DR EMBL; AP006878; BAD85372.1; -; Genomic_DNA.
DR SMR; Q5JGD1; 3-295.
DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001714; Pept_M24_MAP.
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF00557; Peptidase_M24; 1.
DR PRINTS; PR00599; MAPEPTIDASE.
DR TIGRFAMs; TIGR00501; met_pdase_II; 1.
DR PROSITE; PS01202; MAP_2; 1.
DR AminoPeptidase; Cobalt; Complete proteome; Hydrolase; Protease.
KW Aminopeptidase; Cobalt; Complete proteome; Hydrolase; Protease.
SQ SEQUENCE 295 AA; 33021 MW; 4199B9ADA77AD173 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 41 ELGGKPA 47
| | | | |
Db 44 ELGGKPA 50

RESULT 493
Q62KT6_BURMA PRELIMINARY; PRT; 297 AA.
AC Q62KT6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Transcriptional regulator, LysR family.
GN OrderedLocusNames=EMA0958;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
```

```
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Roming C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
CC -!- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
DR EMBL; CP000010; AAU49415.1; -; Genomic_DNA.
DR TIGR; BMA0958; -;
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR000847; HTH_LysR.
DR InterPro; IPR005119; LysR_subst.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF03466; LysR_substrate; 1.
DR PRINTS; PR01590; HTHFIS.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS09331; HTH_LYSR; 1.
KW Complete proteome; DNA-binding; Transcription;
KW Transcription regulation.
SQ SEQUENCE 297 AA; 31288 MW; 9FE1A9ABF68E18EE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 297;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 LAALAA 27
| | | | |
Db 186 LAALAA 192

RESULT 494
Q5UXA2_HALMA PRELIMINARY; PRT; 298 AA.
ID Q5UXA2;
AC Q5UXA2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Methionine aminopeptidase (EC 3.4.11.18).
GN Name=map; OrderedLocusNames=rrnAC3421;
OS Haloarcula marismortui (Halo bacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Haloarcula.
OX NCBI_TaxID=2238;
RN [1];
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 43049;
RX PubMed=15520287; DOI=10.1101/gr.2700304;
RA Baliga N.S., Bonneau R., Facciotti M.T., Pan M., Glusman G.,
RA Deutsch E.W., Shannon P., Chiu Y., Weng R.S., Gan R.R., Hung P.,
RA Date S.V., Marcotte E., Hood L., Ng W.V.;
RT "Genome sequence of Haloarcula marismortui: a halophilic archaeon from
RT the Dead Sea.";
RL Genome Res. 14:2221-2234(2004).
DR EMBL; AY596297; AAU48101.1; -; Genomic_DNA.
DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
DR GO; GO:0008233; P:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001714; Pept_M24_MAP.
DR InterPro; IPR002468; Peptidase_M24.
DR Pfam; PF00557; Peptidase_M24; 1.
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF00557; Peptidase_M24; 1.
```

DR PRINTS: PR00599; MAPEPTIDASE.
 DR TIGRPMs: TIGR00501; met_pdae II; 1.
 KW Aminopeptidase; Cobalt; Complete proteome; Hydrolase; Protease.
 SQ SEQUENCE 298 AA; 32622 MW; 18387822FE7246AF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 298;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ELGGKPA 47
 |||||
 DB 51 ELGGKPA 57

RESULT 495
 Q63UM4_BURPS
 ID Q63UM4_BURPS PRELIMINARY; PRT; 300 AA.
 AC Q63UM4;
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE Putative LysR-family transcriptional regulator.
 GN OrderedLocuNames=BPS1572;
 OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; pseudomallei group.
 CX NCBI_TaxID=28450;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K96243;
 RX PubMed=1537794; DOI=10.1073/pnas.0403302101;
 RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.-M.,
 RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
 RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
 RA Brooke K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
 RA Chillingworth T., Cronin A., Crossset B., Davis P., Deshazer D.,
 RA Felwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
 RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
 RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
 RA Songvilailai S., Stevens K., Tumapa S., Vesaratchavest M.,
 RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
 RT "Genomic plasticity of the causative agent of melioidosis,
 RT Burkholderia pseudomallei";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
 CC -1- SIMILARITY: Contains 1 HTH LysR-type DNA-binding domain.
 DR EMBL; BX571965; CAH35573.1; -; Genomic_DNA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR InterPro; IPR002197; HTH_Fis.
 DR InterPro; IPR000847; HTH_LysR.
 DR InterPro; IPR005119; LysR_subet.
 DR Pfam; PF01126; HTH 1; 1.
 DR Pfam; PF01456; LysR_substrate; 1.
 DR PRINTS; PR01590; HTHFIS.
 DR PRINTS; PR00039; HTHLYSR.
 DR PROSITE; PS50931; HTH_LYSR; 1.
 KW Complete proteome; DNA-binding; Transcription;
 KW Transcription regulation.
 SQ SEQUENCE 300 AA; 31560 MW; 21EDFC79AD5A7631 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 300;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 27
 |||||
 DB 189 LAALAA 195

RESULT 496
 Q7U8H1_SYNPX
 ID Q7U8H1_SYNPX PRELIMINARY; PRT; 302 AA.

AC Q7U8H1;
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
 DE Putative dTDP-4-dehydrohamose reductase (EC 1.1.1.133).
 GN NameRfbD; OrderedLocuNames=SYNW0647;
 OS Synechococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 CX NCBI_TaxID=84588;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
 RA Palenik B., Brahamsha B., Larimer F.W., Land M.L., Hauser L.,
 RA Chain P., Lamerdin J.E., Regala W., Allen E.B., McCarren J.,
 RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
 RT "The genome of a motile marine Synechococcus";
 RL Nature 424:1037-1042(2003).
 DR EMBL; BX569690; CAE07162.1; -; Genomic_DNA.
 DR HSPSP; P26392; 1KBZ.
 DR GO; GO:0008831; F:dTDP-4-dehydrohamose reductase activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0045226; P:extracellular polysaccharide biosynthesis; IEA.
 DR InterPro; IPR005913; TDP_rham_reduct.
 DR Pfam; PF04321; RmlD_sub_Bind; 1.
 DR TIGRPMs; TIGR01214; rmlD; 1.
 KW Complete proteome; Oxidoreductase.
 SQ SEQUENCE 302 AA; 31836 MW; 30961D90EB24E2C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 302;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQRA 93
 |||||
 DB 242 LGLLQRA 248

RESULT 497
 Q57FK0_BRUAB
 ID Q57FK0_BRUAB PRELIMINARY; PRT; 303 AA.
 AC Q57FK0;
 DT 10-MAY-2005 (T-EMBLrel. 30, Created)
 DT 10-MAY-2005 (T-EMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (T-EMBLrel. 30, Last annotation update)
 DE Tetrapyrrole methylase family protein.
 GN OrderedLocuNames=Brubbi_0173;
 OS Brucella abortus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 CX NCBI_TaxID=235;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=9-941 / Biovar 1;
 RX PubMed=15805518; DOI=10.1126/JB.187.8.2715-2726.2005;
 RA Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuerner R.L.,
 RA Qing Z., Li L.-L., Kapur V., Alt D.P., Olsen S.C.;
 RT "Completion of the genome sequence of Brucella abortus and comparison
 RT to the highly similar genomes of Brucella melitensis and Brucella
 RT suis";
 RL J. Bacteriol. 187:2715-2726(2005).
 DR EMBL; AB017223; AAX73584.1; -; Genomic_DNA.
 KW Complete proteome; Methyltransferase.
 SQ SEQUENCE 303 AA; 32076 MW; 29995734CEC028D9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 303;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 |||||
 DB 139 VLAALAA 145


```

RESULT 498
Q8YEV7 BRUME PRELIMINARY; PRT; 303 AA.
AC Q8YEV7;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DE 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE METHYLTRANSFERASE (EC 2.1.1.-)
GN OrderedLocusNames=BMEI1770;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bruciellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Muej C., Los T.,
RA Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyripides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
DR EMBL; AB009610; AAL52951.1; -; Genomic_DNA.
DR PIR; AD3473; AD3473.
DR GO; GO:0008168; F:methyltransferase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR008778; Cor/por_Mettransf.
DR InterPro; IPR008189; UPF0011.
DR Pfam; PF00590; TP_methylase; 1.
DR TIGRFAMs; TIGR005917; Mase_Yral; 1.
DR TIGRFAMs; TIGR00096; UPF0011; 1.
DR PROSITE; PS01296; UPF0011; 1.
KW Complete proteome.
SQ SEQUENCE 303 AA; 32077 MW; 29955734CEC028D9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 303;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 139 VLAALAA 145

RESULT 499
ISPE MYCBO STANDARD; PRT; 306 AA.
AC P65179; O05596;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148) (CMK)
DE (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase).
GN Name=ispE; OrderedLocusNames=Mb1038;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12789872; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- FUNCTION: Catalyzes the phosphorylation of the position 2 hydroxy

group of 4-diphosphocytidyl-2C-methyl-D-erythritol (By
similarity).
-!- CATALYTIC ACTIVITY: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-
erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-
D-erythritol.
-!- PATHWAY: Isoprenoid biosynthesis; isopentenyl-PP biosynthesis via
DXP pathway; isopentenyl-PP from 1-deoxy-D-xylulose 5-phosphate;
step 3.
-!- SIMILARITY: Belongs to the GMP kinase family. IspE subfamily.
-----
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
-----
EMBL; BX248337; CAD93899.1; -; Genomic_DNA.
HMAP; MF_00061; -; 1.
InterPro; IPR006204; GMP_kinase.
InterPro; IPR004424; IspE.
Pfam; PF00288; GMP_kinases; 1.
DR PIRSF; PIRSF010376; IspE; 1.
DR TIGRFAMs; TIGR00154; ispE; 1.
KW ATP-binding; Complete proteome; Isoprene biosynthesis; Kinase;
Nucleotide-binding; Transferase.
FT NP BIND 98 108 ATP (Potential).
FT ACT_SITE 13 13 By similarity.
FT ACT_SITE 140 140 By similarity.
FT SEQUENCE 306 AA; 31382 MW; 98A513F46E487492 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 206 VLAALAA 212

RESULT 500
ISPE MYCTU STANDARD; PRT; 306 AA.
AC P65178; O05596;
DT 30-MAY-2000 (Rel. 39, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148) (CMK)
DE (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase).
GN Name=ispE; OrderedLocusNames=Rv1011, MT1040; ORFNames=MTIC237.28;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;

```

```

RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey B.K., Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
CC -1- FUNCTION: Catalyzes the phosphorylation of the position 2 hydroxy
CC group of 4-diphosphocytidylyl-2C-methyl-D-erythritol (By
CC similarity).
CC -1- CATALYTIC ACTIVITY: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-
CC erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-
CC D-erythritol.
CC -1- PATHWAY: Isoprenoid biosynthesis; isopentenyl-PP biosynthesis via
CC DXF pathway; isopentenyl-PP from 1-deoxy-D-xylulose 5-phosphate:
CC step 3.
CC -1- SIMILARITY: Belongs to the GHMP kinase family. IspE subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX842575; CAB08138.1; -; Genomic DNA.
CC EMBL; AB000516; AAK45290.1; -; Genomic DNA.
CC PIR; F70603; F70603.
CC TIGR; MT1040; -.
CC TuberculList; Rv1011; -.
CC HAMAP; MF_00061; -.
CC InterPro; IPR006204; GHMP_kinase.
CC InterPro; IPR004424; IspE.
CC Pfam; PF00288; GHMP_kinases; 1.
CC PIRSF; PIRSF010376; IspE; 1.
CC TIGRFAMs; TIGR00154; ispE; 1.
CC ATP-binding; Complete proteome; Isoprene biosynthesis; Kinase;
KW Nucleotide-binding; Transferase.
FT NP_BIND 98 108 ATP (Potential).
FT ACT_SITE 13 13 By similarity.
FT ACT_SITE 140 140 By similarity.
SQ SEQUENCE 306 AA; 31382 MW; 98A513F46E8487492 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 206 VLAALAA 212
|||||
|||||

RESULT 501
OTC_AC1AD STANDARD; PRT; 306 AA.
AC Q6FC8;
DT 01-FEB-2005 (Rel. 46, Created)
DT 01-FEB-2005 (Rel. 46, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Ornithine carbamoyltransferase (EC 2.1.3.3) (OTCase).
GN Name=argP; OrderedLocNames=AC1AD1366;
OS Acinetobacter sp. (strain ADP1).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=62977;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15514110; DOI=10.1093/nar/gkn910;
RA Barbe V., Vallet D., Fonknechten N., Kreimeyer A., Oztas S.,
RA Labarre L., Cruveiller S., Robert C., Duprat S., Wincker P.,
RA Orneton L.N., Weissenbach J., Marliere P., Cohen G.N., Medigue C.;

```

```

RT "Unique features revealed by the genome sequence of Acinetobacter sp.
RT ADP1, a versatile and naturally transformation competent bacterium.";
RL Nucleic Acids Res. 32:5766-5779(2004).
CC -1- CATALYTIC ACTIVITY: Carbamoyl phosphate + L-ornithine = phosphate
CC + L-citrulline.
CC -1- PATHWAY: Arginine biosynthesis; sixth step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the ATCase/OTCase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; CR543861; CAG68233.1; -; Genomic DNA.
CC HAMAP; MF_01109; -.
CC InterPro; IPR006130; Asp/Orn_COTranf.
CC InterPro; IPR006131; Asp_Orn_bd.
CC InterPro; IPR006132; OCTace_F_bd.
CC InterPro; IPR002292; Orn_carbamitrans.
CC Pfam; PF00185; OCTace; 1.
CC Pfam; PF02729; OCTace_N; 1.
CC PIRSF; PIRSF000416; OCT_ATCase; 1.
CC PRINTS; PR00100; ACTCASE.
CC PRINTS; PR00102; OTCase.
CC TIGRFAMs; TIGR00658; orn_carb_tr; 1.
CC PROSITE; PS00097; CARBAMOYLTRANSFERASE; 1.
KW Amino-acid biosynthesis; Arginine biosynthesis; Complete proteome;
KW Transferase.
FT SITE 28 28 Important for structural integrity (By
FT SITE 53 57 similarity).
FT SITE 104 104 Carbamoylphosphate binding (By
FT SITE 131 131 similarity).
FT SITE 144 144 Carbamoylphosphate binding (By
FT SITE 258 261 similarity).
FT SITE 306 AA; 34468 MW; C44AF7CB25DD32E7 CRC64;
SQ SEQUENCE 306 AA; 34468 MW; C44AF7CB25DD32E7 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 306;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 82 FKGVKVLG 88
Db 40 FKGVKVLG 46
|||||
|||||

RESULT 502
Q9YCE0_AERPE
ID Q9YCE0_AERPE PRELIMINARY; PRT; 306 AA.
AC Q9YCE0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 306aa long hypothetical methanol dehydrogenase regulator.
GN OrderedLocNames=APE1317;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STAIN-K1;
RX MEDLINE=9310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankael A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,

```

RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RT Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RL crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; BA000002; BAA80308.1; -; Genomic_DNA.
DR PIR; F72606; F72606
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR011703; AAA 3.
DR InterPro; IPR003593; AAA_ATPase.
DR Pfam; PF07726; AAA3; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding; Complete proteome; Hypothetical protein;
KW Nucleotide-binding; Transport.
SQ SEQUENCE 306 AA; 33720 MW; 850278C5A42304BB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 VLAALAA 26
| | | | |
Db 26 VLAALAA 32

RESULT 503
Q4HCY3_9DEIO
ID Q4HCY3_9DEIO PRELIMINARY; PRT; 306 AA.
AC Q4HCY3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Binding-protein-dependent transport systems inner membrane
DE component.
GN ORFNames=DgeODRAFT_1921;
OS Deinococcus geothermalms DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hamon N., Israni S., Pittluck S., Richardson P.;
RA Hamon N., Israni S., Lapidus A., Barry K., Dettler C., Glavina T.,
RT "Sequencing of the draft genome assembly of Deinococcus geothermalms
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Deinococcus geothermalms
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
DR EMBL; AAHE01000001; EAL03815.1; -; Genomic DNA.
SQ SEQUENCE 306 AA; 33847 MW; 5A09AEEBAE3E4491 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 306;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 21 LAALAA 27
| | | | |
Db 110 LAALAA 116

RESULT 504
Q5GV13_XANOR
ID Q5GV13_XANOR PRELIMINARY; PRT; 307 AA.
AC Q5GV13;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=X003154;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gki206;
RA Lee B.-W., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice.";
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AB013598; AAW76408.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 307 AA; 34114 MW; 7F28C7D41AEFF54 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 307;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 VLAALAA 26
| | | | |
Db 284 VLAALAA 290

RESULT 505
Q6IQ36_CAEBR
ID Q6IQ36_CAEBR PRELIMINARY; PRT; 310 AA.
AC Q6IQ36;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG07194 (Fragment).
GN Name=CBG07194;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.Briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
DR EMBL; CAAC01000031; CAB62976.1; -; Genomic_DNA.
DR GO; GO:0016030; C:membrane; IEA.
DR GO; GO:0004930; F:G-protein coupled receptor activity; IEA.
DR InterPro; IPR003002; 7TM chemrecept1.
DR InterPro; IPR000168; Nm7TM_chemrecept.
DR Pfam; PF01461; 7tm4; 1.
KW Hypothetical protein.
FT NON TER 1
FT NON TER 310 310
SQ SEQUENCE 310 AA; 35849 MW; 62E8225E8C5553BD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 310;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 SVGCWVI 36
 |||||
 Db 198 SVGCWVI 204

RESULT 506

Q5YWF0_NOCFA
 ID Q5YWF0_NOCFA PRELIMINARY; PRT; 310 AA.
 AC Q5YWF0;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative membrane protein.
 GN OrderedLocusNames=afa26440;
 OS Nocardia farcinica.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Nocardiaceae; Nocardia.
 OC NCBI_TaxID=37329;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IFW 10152;
 RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
 RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
 RA Shiba T., Hattori M.;
 RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
 DR EMBL; AF006618; BA057491.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR000620; DUF6_TM.
 DR Pfam; PF00892; DUF6; 2.
 KW Complete proteome.
 SQ SEQUENCE 310 AA; 31866 MW; 44AB9AF9B62391A7 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 310;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 |||||
 Db 279 VLAALAA 285

RESULT 507

O33744_STRSQ
 ID O33744_STRSQ PRELIMINARY; PRT; 312 AA.
 AC O33744;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein strB.
 GN Name=stb;
 OS Streptomyces sp.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1931;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=F20;
 RX MEDLINE=98037487; PubMed=9371436;
 RA Fernandez-Moreno M.A., Vallin C., Maltapida F.;
 RT "Streptothricin biosynthesis is catalyzed by enzymes related to
 RT nonribosomal peptide bond formation.";
 RL J. Bacteriol. 179:6929-6936(1997).
 DR EMBL; Y10293; CAA71337.1; -; Genomic_DNA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR InterPro; IPR02575; APH_trans.
 DR Pfam; PF01636; APH; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 312 AA; 33832 MW; 780A654F543B4B9F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 27
 |||||
 Db 18 LAALAA 24

RESULT 508

Q4J105_AZOV1
 ID Q4J105_AZOV1 PRELIMINARY; PRT; 312 AA.
 AC Q4J105;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Dehydrogenase precursor.
 GN ORFNames=AVINDRAFT_4883;
 OS Azotobacter vinelandii AVOP.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Azotobacter.
 OC NCBI_TaxID=322710;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AVOP;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Azotobacter vinelandii
 RT AVOP.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AVOP;
 RG US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Azotobacter vinelandii
 RT AVOP.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AVOP;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAAU03000002; EAM06522.1; -; Genomic_DNA.
 KW Signal.
 FT SIGNAL 312 312 Potential.
 SQ SEQUENCE 312 AA; 31735 MW; 57E04B13833E2B17 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
 |||||
 Db 78 GGVLAAL 84

RESULT 509

Q4K150_PSEF5
 ID Q4K150_PSEF5 PRELIMINARY; PRT; 313 AA.
 AC Q4K150;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Transcriptional regulator, MerR family.

Query Match 5.9%; Score 7; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GN ORFNames=PFL_0952;
OC Pseudomonas fluorescens (strain Pf-5).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=220664;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PF-5;
RX PubMed=15980861; DOI=10.1038/nbt1110;
RA Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA Mavrodi D., DeBoy R.T., Seshadri R., Ren O., Madupu R., Dodson R.J.,
RA Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA Winn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA Khouri H.M., Pierson E., Pierson L. III, Thomasow L., Loper J.,
RT "Complete genome sequence of the plant commensal Pseudomonas
RT fluorescens Pf-5.";
RT Nat. Biotechnol. 23:873-878(2005).
RL Nat. Biotechnol. 23:873-878(2005).
DR EMBL: CP000076; AAY90239.1; -; Genomic DNA.
SQ SEQUENCE 313 AA; 34834 MW; B198DB9389DBBE8D CRC64;

Query Match 5.9%; Score 7; DB 2; Length 313;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRATQ 95
Db |||||
187 LLQRATQ 193

RESULT 510
TRMB_PROAC STANDARD; PRT; 314 AA.
AC Q6A6S9;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE tRNA (Guanine-N(7)-methyltransferase (EC 2.1.1.33) (tRNA(m7G46)-
DE methyltransferase).
GN Name=trmB; OrderedLocusNames=PPA1808;
OS Propionibacterium acnes.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Propionibacteriaceae; Propionibacteriidae; Propionibacterium.
OX NCBI_TaxID=1747;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KPA171202 / DSM 16379;
RX PubMed=15286373; DOI=10.1126/science.1100330;
RA Brueggemann H., Henne A., Hoster F., Liebegang H., Wierer A.,
RA Strittmatter A., Hujer S., Duerre P., Gottschalk G.;
RT "The complete genome sequence of Propionibacterium acnes, a commensal
RT of human skin.";
RL Science 305:671-673(2004).
CC -!- FUNCTION: Catalyzes the formation of N(7)-methylguanine at
CC position 46 (m7G46) in tRNA (by similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC homocysteine + tRNA containing N(7)-methylguanine.
CC -!- SIMILARITY: Belongs to the RNA m7G methyltransferase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL: AB017283; AAT93534.1; -; Genomic DNA.
DR HAMAP; MF_01057; -; 1.
DR InterPro; IPR004395; CHP91.
DR InterPro; IPR003358; Methyltransf_4.
DR InterPro; IPR000051; SAM bd.
DR PANTHER; PTHR12793; Methyltransf_4; 1.
DR Pfam; PF03390; Methyltransf_4; 1.
DR TIGRFAMs; TIGR00091; Cons Hypoth91; 1.
KW Complete proteome; Methyltransferase; Transferase; tRNA processing.

SQ SEQUENCE 314 AA; 34701 MW; B5C1BE1975F4C9DB CRC64;

Query Match 5.9%; Score 7; DB 1; Length 314;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db |||||
149 VLAALAA 155

RESULT 511
Q89L42 BRAJA PRELIMINARY; PRT; 315 AA.
AC Q89L42;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B114706 protein.
GN OrderedLocusNames=b114706;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
[1]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL: BA000040; BAC49971.1; -; Genomic DNA.
DR HSSP; Q50245; 110Q.
DR InterPro; IPR011964; Beta_rpt_yvtn.
DR TIGRFAMs; TIGR02276; beta_rpt_yvtn; 3.
KW Complete proteome.
SQ SEQUENCE 315 AA; 32363 MW; F38C7E6956AC0233 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db |||||
5 VLAALAA 11

RESULT 512
ISPE MYCPA STANDARD; PRT; 316 AA.
ID ISPE MYCPA
AC Q741WI;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (EC 2.7.1.148) (CMK)
DE (4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase).
GN Name=ispE; OrderedLocusNames=MAP0976;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=k10;
RA Li L., Bamanine J., Zhang Q., Anonin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Catalyzes the phosphorylation of the position 2 hydroxy
CC group of 4-diphosphocytidyl-2C-methyl-D-erythritol (By

CC similarity).

CC -1- CATALYTIC ACTIVITY: ATP + 4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol = ADP + 2-phospho-4-(cytidine 5'-diphospho)-2-C-methyl-D-erythritol.

CC -1- PATHWAY: Isoprenoid biosynthesis; isopentenyl-PP biosynthesis via DXP pathway; isopentenyl-PP from 1-deoxy-D-xylulose 5-phosphate: step 3.

CC -1- SIMILARITY: Belongs to the GHMP kinase family. IspE subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; AB017230; AA030293.1; -; Genomic_DNA.

DR HAMAP; MF 00061; -; 1.

DR InterPro; IPR006204; GHMP_kinase.

DR InterPro; IPR004424; IspE.

DR Pfam; PF00288; GHMP_kinase; 1.

DR PIRSF; PIRSF010376; IspE; 1.

DR TIGRFAMs; TIGR00154; IspE; 1.

DR ATP-binding; Complete proteome; Isoprene biosynthesis; Kinase; Nucleotide-binding; Transferase.

KW NP_BIND 108 118 ATP (Potential).

FT ACT_SITE 23 23 By similarity.

FT ACT_SITE 150 150 By similarity.

FT ACT_SITE 316 AA; 32460 MW; 9DB2627B6C9CAFA7 CRC64;

SQ SEQUENCE 316 AA; 5.9%; Score 7; DB 1; Length 316;

Query Match Best Local Similarity 100.0%; Pred. No. 4.1e+02; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26

Db 216 VLAALAA 222

|||||

|||||

|||||

|||||

RESULT 513

Q9VUNI DROME

ID Q9VUNI DROME PRELIMINARY; PRT; 316 AA.

AC Q9VUNI;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE CG13698-PB (GH19567p).

GN ORFNames=CG13698;

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OC NCBI_TaxID=7227;

OX NCBI_TaxID=7227;

RN [1]

RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gockney J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Vardell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischman W., Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., Palazzolo R., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shan H., Shue B.C., Sinden-Klamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

*The genome sequence of *Drosophila melanogaster*.*;

RL Science 287:2185-2195(2000).

RL [2]

RP NUCLEOTIDE SEQUENCE.

RP MEDLINE=22426065; PubMed=12537568;

RA Celinker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A., George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R., Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J., Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;

*Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence.*;

RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).

RL [3]

RP NUCLEOTIDE SEQUENCE.

RP MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svirskaas R., Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M., Ashburner M., Celinker S.E.;

*The transposable elements of the *Drosophila melanogaster* euchromatin: a genomics perspective.*;

RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).

RL [4]

RP NUCLEOTIDE SEQUENCE.

RP MEDLINE=22426069; PubMed=12537572;

RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;

*Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review.*;

RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).

RL [5]

RP NUCLEOTIDE SEQUENCE.

RP Berkeley Drosophila Genome Project;

RA Celinker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R., Hoskins R., Stapleton M., Pacleb J., Park S., Svirskaas R., Smith E., Yu C., Rubin G.;

Drosophila melanogaster release 4 sequence.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.

RL [6]

RP NUCLEOTIDE SEQUENCE.

RP FlyBase;

RL Submitted (MAR-2005) to the EMBL/GenBank/DBSJ databases.

RL [7]

RP NUCLEOTIDE SEQUENCE.

RP STRAIN=Berkeley;

RC Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,

```

RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003521; AAF49280.1; -, Genomic_DNA.
DR EMBL; AY051547; AAK92971.1; -, mRNA.
DR Ensembl; CG13698; Drosophila melanogaster.
DR FlyBase; FBgn0036773; CG13698.
SQ SEQUENCE 316 AA; 33527 MW; 285B3B0F5983D604 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 316;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 127 LLGGVLA 133

RESULT 514
Q5TBB3 HUMAN
ID Q5TBB3_HUMAN PRELIMINARY; PRT; 319 AA.
AC Q5TBB3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Solute carrier family 9 (Sodium/hydrogen exchanger), isoform 1
DE (Antipporter, Na+/H+, aniloride sensitive).
GN Name=SLC9A1; ORFNames=RP4-633N17.1-003;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137860; CA122086.1; -, Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0015385; F:sodium:hydrogen antiporter activity; IEA.
DR GO; GO:0006885; P:regulation of pH; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001970; Naf_exchng_r1.
DR PRINTS; PR01085; NAHEXCHNGRI.
SQ SEQUENCE 319 AA; 36164 MW; 15C9E75F9D0691AC CRC64;

Query Match          5.9%; Score 7; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLGL 89
Db 242 KGKVLGL 248

RESULT 515
Q72429 HUMAN
ID Q72429_HUMAN PRELIMINARY; PRT; 319 AA.
AC Q72429;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Lambda-crystallin (Crystallin, lambda 1).
GN Name=CRYL1; ORFNames=RP11-187L3.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Xu Z.G., Yu L., Bi A.D., Cui W.C., Huang J., Zhao S.Y.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
```

```

RP NUCLEOTIDE SEQUENCE.
RA Yang J., Yu L., Dai F.Y., Cui W.C., Zheng L.H., Zhao S.Y.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Tromans A.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087898; AAP97197.1; -, mRNA.
DR EMBL; AL161715; CAH71074.1; -, Genomic_DNA.
DR EMBL; AL590096; CAH71074.1; -, Genomic_DNA.
DR EMBL; AL590096; CAH71074.1; JOINED; Genomic_DNA.
DR EMBL; AL161715; CAH71074.1; JOINED; Genomic_DNA.
DR Ensembl; ENSG00000165475; Homo sapiens.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006631; P:fatty acid metabolism; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR InterPro; IPR006180; 3HCDH.
DR InterPro; IPR006108; 3HCDH_C.
DR InterPro; IPR006176; 3HCDH_NAD_bd.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF00725; 3HCDH_1.
DR Pfam; PF02737; 3HCDH_N; 1.
DR PROSITE; PS00067; 3HCDH; UNKNOWN 1.
SQ SEQUENCE 319 AA; 35419 MW; C8DCF74DDE2FDE21 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 GCVVIVG 38
Db 7 GCVVIVG 13

RESULT 516
Q5RDZ2 PONPY
ID Q5RDZ2_PONPY PRELIMINARY; PRT; 319 AA.
AC Q5RDZ2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein DRP2p469F1537.
GN Name=DKF2p469F1537;
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUR=Kidney;
RG The German cDNA Consortium;
RA Ottenwaelder B., Obermaier B., Deutschenbaur S., Schaiipp A.,
RA Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;
RL Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR857749; CAH90015.1; -, mRNA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006631; P:fatty acid metabolism; IEA.
DR GO; GO:0006629; P:lipid metabolism; IEA.
DR InterPro; IPR006180; 3HCDH.
DR InterPro; IPR006108; 3HCDH_C.
DR InterPro; IPR006176; 3HCDH_NAD_bd.
DR InterPro; IPR000205; NAD_BS.
DR Pfam; PF00725; 3HCDH_1.
DR Pfam; PF02737; 3HCDH_N; 1.
DR PROSITE; PS00067; 3HCDH; UNKNOWN 1.
KW Hypothetical protein.
SQ SEQUENCE 319 AA; 35437 MW; 641CED57DA2BDE34 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 32 GCVVIVG 38
Db 7 GCVVIVG 13

RESULT 517
Q727G8_DESVH
ID Q727G8_DESVH PRELIMINARY; PRT; 322 AA.
AC Q727G8; 322 AA.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Membrane protein, putative.
GN OrderedLocusNames=DVU2887;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Sehadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017318; AAS97359.1; -; Genomic_DNA.
DR TIGR; DVU2887; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6_TM.
DR Pfam; PF00892; DUF6; 2.
KW Complete proteome.
SQ SEQUENCE 322 AA; 33267 MW; A7DA561F1AC55A6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 37 VLAALAA 43

RESULT 518
Q72DA2_DESVH
ID Q72DA2_DESVH PRELIMINARY; PRT; 323 AA.
AC Q72DA2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DVU1027;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Sehadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;

RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017312; AAS95507.1; -; Genomic_DNA.
DR TIGR; DVU1027; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 323 AA; 34363 MW; 6840204E43FA2A1 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 323;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
Db 137 VLLGGVL 143

RESULT 519
Q4TN16_9SPHN
ID Q4TN16_9SPHN PRELIMINARY; PRT; 327 AA.
AC Q4TN16;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Membrane protein.
GN ORFNames=EL11761;
OS Erythrobacter litoralis HTCC2594.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC Sphingomonadaceae; Erythrobacter.
OX NCBI_TaxID=314225;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HTCC2594;
RA Giovannoni S.J., Cho J.-C., Ferriera S., Johnson J., Kravitz S.,
RA Halpern A., Remington K., Besson K., Tran B., Rogers Y.-H.,
RA Friedman R., Venter J.C.;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAGG01000005; EAL75250.1; -; Genomic DNA.
SQ SEQUENCE 327 AA; 33352 MW; A5506EE68DC9D1B1 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 327;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 65 VLAALAA 71

RESULT 520
Q89R48_BRAJA
ID Q89R48_BRAJA PRELIMINARY; PRT; 328 AA.
AC Q89R48;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Amino acid ABC transporter permease protein.
GN OrderedLocusNames=blr2924;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpō S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;

RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 RL Bradyrhizobium japonicum USDA110.";

DR EMBL; BA000040; BAC48189.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transpore; IEA.
 DR InterPro; IPR001851; Bac_inmem_transp.
 DR Pfam; PF02653; BPD_transp_2; 1.
 KW Complete proteome.

SQ SEQUENCE 328 AA; 35805 MW; 7CB69127DF31E346 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 328;
 Best Local Similarity 100.0%; Pred. No. 4.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 19 VLAALAA 25

RESULT 521

Q5TY15 ANOQA PRELIMINARY; PRT; 331 AA.
 AC Q5TY15;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE ENGANGP0000028488.
 GN ORFNames=ENSANGG00000023215;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
 OC Anophelinae; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RT "Anopheles gambiae re-annotation.";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; AAB01004767; EAL42269.1; -; Genomic_DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR InterPro; IPR000160; GDEF.
 DR InterPro; IPR004837; NaCa_Exmemb.
 DR InterPro; IPR002040; Tachy_Neurokinin.
 DR Pfam; PF00990; GDEF; 1.
 DR Pfam; PF01699; Na_Ca_ex; 1.
 DR SMART; SM00267; DUF1; 1.
 DR TIGRFAMs; TIGR00254; GDEF; 1.
 DR PROSITE; PS00887; GDEF; 1.
 DR PROSITE; PS00267; TACHYKININ; UNKNOWN 1.
 DR PROSITE; PS00267; TACHYKININ; UNKNOWN 1.
 SQ SEQUENCE 331 AA; 34351 MW; 339E768D78DD7D3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 331;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 41 ELGGKPA 47
 |||||
 Db 315 ELGGKPA 321

RESULT 522

Q82PF3_STRAW

ID Q82PF3_STRAW PRELIMINARY; PRT; 332 AA.
 AC Q82PF3;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative ribose/xylose/arabinose/galactoside ABC transporter permease
 DE protein.
 GN OrderedLocNames=SAV968;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RA "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis.";
 RL Nat. Biotechnol. 21:526-531(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL; BA000030; C:membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transpore; IEA.
 DR InterPro; IPR001851; Bac_inmem_transp.
 DR Pfam; PF02653; BPD_transp_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 332 AA; 33754 MW; E40A52057F5CC9B9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 332;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
 |||||
 Db 89 LGGVLA 95

RESULT 523

Q60GF1_COMAC PRELIMINARY; PRT; 335 AA.
 ID Q60GF1_COMAC PRELIMINARY;
 AC Q60GF1;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Reductase.
 GN Name=ORF7NE;
 OS Comamonas acidovorans (Pseudomonas acidovorans).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Comamonadaceae; Delftia.
 OX NCBI_TaxID=80866;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=7N;
 RX PubMed=15618615; DOI=10.1271/bbb.68.2457;
 RA Urata M., Uchida E., Nojiri H., Otori T., Obo R., Miyamura N.,
 RA Ouchiya N.;
 RT "Genes Involved in Aniline Degradation by Delftia acidovorans Strain
 RT 7N and Its Distribution in the Natural Environment.";
 RL BioSci. Biotechnol. Biochem. 68:2457-2465(2004).
 DR EMBL; AB177545; BAD61051.1; -; Genomic_DNA.

DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR006058; 2Fe2S fd BS.
DR InterPro; IPR001834; Cyt B5 reductase.
DR InterPro; IPR008333; FAD_binding_6.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR001709; FPN_Cyt_reductase.
DR InterPro; IPR001433; Oxred FAD/NAD(P).
DR Pfam; PF00970; FAD_binding_6; 1.
DR Pfam; PF00111; Fer2; 1.
DR Pfam; PF00175; NAD_binding_1; 1.
DR PRINTS; PR00406; CYTB5RDTASE.
DR PRINTS; PR00371; FPNCR.
DR PRINTS; PR00409; PHDIOXRDTASE.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; UNKNOWN 1.
SQ SEQUENCE 335 AA; 35598 MW; A72ED3EP262289DD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 211 VLAALAA 217

RESULT 524
Q6TFU0_9PSBD
ID Q6TFU0_9PSBD PRELIMINARY; PRT; 335 AA.
AC Q6TFU0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Aniline dioxygenase reductase.
OS Pseudomonas sp. K82.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=251433;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K82; Kim J.-Y., Kim E.-A.;
RA Kim S.-I., Kim J.-Y., Kim E.-A.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY422718; AAR03448.1; -; Genomic DNA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016702; F:oxidoreductase activity, acting on single d. .; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR006058; 2Fe2S fd BS.
DR InterPro; IPR001834; Cyt B5 reductase.
DR InterPro; IPR008333; FAD_binding_6.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR001709; FPN_Cyt_reductase.
DR InterPro; IPR001433; Oxred FAD/NAD(P).
DR InterPro; IPR00951; Phdiox reductase.
DR Pfam; PF00970; FAD_binding_6; 1.
DR Pfam; PF00111; Fer2; 1.
DR Pfam; PF00175; NAD_binding_1; 1.
DR PRINTS; PR00406; CYTB5RDTASE.
DR PRINTS; PR00371; FPNCR.
DR PRINTS; PR00409; PHDIOXRDTASE.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; UNKNOWN 1.
KW Dioxygenase.
SQ SEQUENCE 335 AA; 35698 MW; A32CD8F3312991C8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 211 VLAALAA 217

RESULT 525
Q9EXM3_PSESP
ID Q9EXM3_PSESP PRELIMINARY; PRT; 335 AA.
AC Q9EXM3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ORP2.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Y-2;
RX MEDLINE=98276889; PubMed=9614705;
RA Murakami S., Nakanishi Y., Kodama N., Takenaka S., Shinke R., Aoki K.;
RT "Purification, characterization, and gene analysis of catechol 2,3-
dioxygenase from the aniline-assimilating Pseudomonas species AW-2.";
RL Biosci. Biotechnol. Biochem. 62:747-752(1998).
DR EMBL; AB004065; BAB18930.1; -; Genomic DNA.
DR HSPSP; P06543; ICZP.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; F:electron transport; IEA.
DR InterPro; IPR006058; 2Fe2S fd BS.
DR InterPro; IPR001834; Cyt B5 reductase.
DR InterPro; IPR008333; FAD_binding_6.
DR InterPro; IPR001041; Ferredoxin.
DR InterPro; IPR001709; FPN_Cyt_reductase.
DR InterPro; IPR001433; Oxred FAD/NAD(P).
DR InterPro; IPR00951; Phdiox reductase.
DR Pfam; PF00970; FAD_binding_6; 1.
DR Pfam; PF00111; Fer2; 1.
DR Pfam; PF00175; NAD_binding_1; 1.
DR PRINTS; PR00406; CYTB5RDTASE.
DR PRINTS; PR00371; FPNCR.
DR PRINTS; PR00409; PHDIOXRDTASE.
DR PROSITE; PS00197; 2FE2S_FERREDOXIN; UNKNOWN 1.
SQ SEQUENCE 335 AA; 35631 MW; B0A52CC22996CDDC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 211 VLAALAA 217

RESULT 526
Q55N25_CRYNE
ID Q55N25_CRYNE PRELIMINARY; PRT; 339 AA.
AC Q55N25;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=CNBH2590;
OS Cryptococcus neoformans var. neoformans B-3501A.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=283643;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
Wicks B.L., Fu J., Davis R.W.;

RT "Cryptococcus neoformans serotype D sequencing.";
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AA01000042; EAL19160.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 339 AA; 35708 MW; 793A19A97D20B2CA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 339;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGKVLGL 89
 |||||
 Db 320 KGKVLGL 326

RESULT 527
 Q5KBF4 CRYNE PRELIMINARY; PRT; 339 AA.
 AC Q5KBF4;
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=CN102730;
 OS Cryptococcus neoformans var. neoformans JEC21.
 CC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 CC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=214684;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=JEC21;
 RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,
 RA Van Aken S., Fraser C.; to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=JEC21;
 RX PubMed=15653466; DOI=10.1126/science.1103773;
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
 RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
 RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
 RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
 RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinski M.I.,
 RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
 RA Mathewson C.A., Mitchell T.G., Pertea M., Riggs F.R., Salzberg S.L.,
 RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
 RA Suh B.B., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
 RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
 RA Fraser C.M., Hyman R.W.;
 RT "The genome of the basidiomycetous yeast and human pathogen
 RT Cryptococcus neoformans.";
 RL Science 307:1321-1324 (2005).
 DR EMBL; AB017349; AAW45371.1; -; Genomic_DNA.
 DR InterPro; IPR000794; Ketaacyl synth.
 DR InterPro; IPR000408; Reg chr condens.
 DR PRINTS; PR00633; RCNDNSACYN.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
 DR PROSITE; PS00012; RCL_3; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 339 AA; 35681 MW; 4APA2DC1A8A989AE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 339;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGKVLGL 89
 |||||
 Db 320 KGKVLGL 326

RESULT 528
 Q72TU1 LEPIC PRELIMINARY; PRT; 340 AA.
 AC Q72TU1;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Metallo-beta-lactamase.
 GN OrderedLocusNames=LIC10925;
 OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar
 OS Copenhageni)
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=44275;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Fiocruz L1-130;
 RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
 RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B.,
 RA Ho P.L., Haake D.A., Verjovski-Almeida S., Hartskeerl R.A.,
 RA Marques M.V., Oliveira M.C., Menck C.F.M., Leite L.C.C., Carrer H.,
 RA Coutinho L.L., Degraive W.M., Dellagostin O.A., El-Dorfy H.,
 RA Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A.,
 RA Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R.,
 RA Jeronimo S.M.B., Junqueira-de-Azevedo J.L.M., Kimura E.T.,
 RA Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R.,
 RA de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A.,
 RA Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A.,
 RA Camargo L.E.A., Kitajima J.P., Setubal J.C., Van Sluys M.A.;
 RT "Comparative genomics of two Leptospira interrogans serovars reveals
 RT novel insights into physiology and pathogenesis.";
 RT J. Bacteriol. 186:2164-2172 (2004).
 DR EMBL; AB017290; AAS69537.1; -; Genomic DNA.
 GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR001279; Blactamase-like.
 DR Pfam; PF00753; Lactamase_B; 1.
 DR Complete proteome.
 SQ SEQUENCE 340 AA; 38953 MW; 6240027477BAA40 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 340;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 LGLLQRA 93
 |||||
 Db 53 LGLLQRA 59

RESULT 529
 Q8FID2 LEPIN PRELIMINARY; PRT; 340 AA.
 AC Q8FID2;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Metallo-beta-lactamase superfamily protein.
 GN OrderedLocusNames=LA3205;
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
 RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
 RA Zhang Y.-X., Xiong H., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
 RA Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
 RA Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
 RA Yao Z.-J., Shen Y., Qiang B.-O., Xia Q.-C., Guo X.-K., Danchin A.,
 RA Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
 RA Xu J.-G., Zhao G.-P.;
 RT "Unique physiological and pathogenic features of Leptospira
 RT interrogans revealed by whole-genome sequencing.";
 RL Nature 422:888-893 (2003).

```

DR EMBL; AE011482; AAN50402.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR001279; Lactamase-like.
DR Pfam; PF00753; Lactamase_B; 1.
KW Complete proteome.
SQ SEQUENCE 340 AA; 38953 MW; 6240027477BAA40 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 87 LGLLQRA 93
Db 53 LGLLQRA 59
|||||

RESULT 530
Y0401.HUMAN
ID Y0401.HUMAN STANDARD; PRT; 344 AA.
AC Q43151;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical protein KIAA0401 (Pragmat).
GN Name=KIAA0401;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Brain;
RX MEDLINE=38116655; PubMed=9455477;
RA Ishikawa K.-I., Nagase T., Nakajima D., Seki N., Ohira M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro."
RL DNA Res. 4:307-313(1997).
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AB007861; BAA23697.1; -; mRNA.
CC Ensembl; ENSG00000187605; Homo sapiens.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 344 AA; 36962 MW; CC4139DC054BEEB2 CRC64;

Query Match          5.9%; Score 7; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 PAIVPDK 52
Db 130 PAIVPDK 136
|||||

RESULT 531
Q7XT06.ORYSA
ID Q7XT06.ORYSA PRELIMINARY; PRT; 345 AA.
AC Q7XT06;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE OSJNB0050003.16 protein.
GN Name=OSJNB0050003.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Shihartoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12447439; DOI=10.1038/nature01183;
RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
RA Wang Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu Z., Fan D.,
RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
RA Han B.;
RT "Sequence and analysis of rice chromosome 4.";
RL Nature 420:316-320(2002).
DR EMBL; AL606831; CA501726.2; -; Genomic_DNA.
DR Gramene; Q7XT06; -.
DR GO; GO:0004040; F:amidase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000120; Amidase.
DR PANTHER; PTHR11895; Amidase; 1.
DR Pfam; PF01425; Amidase; 1.
DR PROSITE; PS00571; AMIDAS8; 1.
SQ SEQUENCE 345 AA; 36434 MW; 96DD3C8DB68752C0 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 66 VLAALAA 72
|||||

RESULT 532
Q73X82.MYCPA
ID Q73X82.MYCPA PRELIMINARY; PRT; 347 AA.
AC Q73X82;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=NAP2427C;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;
RA Li L., Bannantine J., Zhang Q., Amosin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017235; AAS04744.1; -; Genomic_DNA.
DR InterPro; IPR005321; Peptidase_S58.
DR Pfam; PF03576; Peptidase_S58; 1.
KW Complete proteome.
SQ SEQUENCE 347 AA; 34527 MW; 65D8CC72D665349C CRC64;

Query Match          5.9%; Score 7; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
Db 322 LGGVLA 328
|||||

RESULT 533
PYRD_XYLPA

```

10-0CT-2003 (Rel. 42, Created)
10-0CT-2003 (Rel. 42, Last sequence update)
13-SEP-2005 (Rel. 48, Last annotation update)
Dihydroorotate dehydrogenase (EC 1.3.3.1) (Dihydroorotate oxidase)
(DHODase) (DHODase) (DHODase)
Name=pyrD; OrderedLocNames=Xf2571;
Xylella fastidiosa.
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
Xanthomonadaceae; Xylella.
NCBI_TaxID=2371;
[1]
NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,
Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.D., Gruber A.,
Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
Nhani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,
de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pasquero J.B.,
Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
Vallada H., Van Sluys M.A., Varjovski-Almeida S., Vettore A.L.,
Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
"The genome sequence of the plant pathogen Xylella fastidiosa";
Nature 406:151-159(2000).
-!- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) = orotate +
H(2)O(2).
-!- COPACTOR: Binds 1 FMN per subunit (By similarity).
-!- PATHWAY: Nucleotide biosynthesis; UMP biosynthesis; UMP from
HCO(3)(-); step 4.
-!- SUBUNIT: Homodimer (By similarity).
-!- SUBCELLULAR LOCATION: Inner side of the membrane (By similarity).
-!- SIMILARITY: Belongs to the dihydroorotate dehydrogenase family.
Type 2 subfamily.
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
EMBL; AB004064; AAF85368.1; AUT_INIT; Genomic_DNA.
PIR; P82541; P82541.
HSSP; P05021; 1P76.
HAMAP; MF 00225; -. 1.
InterPro; IPR001295; DHO_dh.
InterPro; IPR005719; DHO_dh2.
InterPro; IPR012135; DHO oxidase.
Pfam; PF01180; DHO_dh; 1.
Pfam; PF01180; DHO oxidase; 1.
TIGRFAMs; TIGR01036; pyrD_sub2; 1.
PROSITE; PS00911; DHODHASE_1; 1.
PROSITE; PS00912; DHODHASE_2; FALSE_NEG.

KW Complete proteome; Flavoprotein; FMN; Oxidoreductase;
KW Pyrimidine biosynthesis.
FT ACT_SITE 175 175 Nucleophile (By similarity).
SQ SEQUENCE 351 AA; 37890 MW; F988D09A2D05F6D8 CRC64;
Query Match 5.9%; Score 7; DB 1; Length 351;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 94 TQQQAVI 100
Db 103 TQQQAVI 109
RESULT 534
PYRD_XYLFT STANDARD; PRT; 351 AA.
AC Q87A77;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dihydroorotate dehydrogenase (EC 1.3.3.1) (Dihydroorotate oxidase)
(DHODase) (DHODase) (DHODase).
GN Name=pyrD; OrderedLocNames=PD1952;
OS Xylella fastidiosa (strain Temeculal / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=193190;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=22421331; PubMed=12533478;
DOI=10.1126/JB.185.3.1018-1026.2003;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.F., da Silva F.R.,
Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
Carrier H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
Kitajima J.P.;
"Comparative analyses of the complete genome sequences of Pierce's
disease and citrus variegated chlorosis strains of Xylella
fastidiosa";
J. Bacteriol. 185:1018-1026(2003).
-!- CATALYTIC ACTIVITY: (S)-dihydroorotate + O(2) = orotate +
H(2)O(2).
-!- COPACTOR: Binds 1 FMN per subunit (By similarity).
-!- PATHWAY: Nucleotide biosynthesis; UMP biosynthesis; UMP from
HCO(3)(-); step 4.
-!- SUBUNIT: Homodimer (By similarity).
-!- SUBCELLULAR LOCATION: Inner side of the membrane (By similarity).
-!- SIMILARITY: Belongs to the dihydroorotate dehydrogenase family.
Type 2 subfamily.
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
EMBL; AB012560; AAO29782.1; -; Genomic_DNA.
HSSP; P05021; 1P76.
HAMAP; MF 00225; -. 1.
InterPro; IPR001295; DHO_dh.
InterPro; IPR005719; DHO_dh2.
InterPro; IPR012135; DHO oxidase.
Pfam; PF01180; DHO_dh; 1.
Pfam; PF01180; DHO oxidase; 1.
TIGRFAMs; TIGR01036; pyrD_sub2; 1.
PROSITE; PS00911; DHODHASE_1; 1.
PROSITE; PS00912; DHODHASE_2; FALSE_NEG.

```
DR PIRSF; PIRSF000164; DHO oxidase; 1.
DR TIGRFAMs; TIGR01036; PYD sub2; 1.
DR PROSITE; PS00911; DHODHASE_2; 1.
DR PROSITE; PS00912; DHODHASE_2; FALSE NEG.
KW Complete proteome; Flavoprotein; FMN; Oxidoreductase;
KW Pyrimidine biosynthesis.
FT ACT SITE 175 175 Nucleophile (By similarity).
SQ SEQUENCE 351 AA; 37877 MW; A3D86E107E2A71C7 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 351;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 TQQAQVI 100
Db 103 TQQAQVI 109
|||||

RESULT 535
Q60KH8 CAEBR
ID Q60KH8 CAEBR PRELIMINARY; PRT; 354 AA.
AC Q60KH8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG24050.
GN Name=CBG24050;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC The C.briggsae Sequencing Consortium;
RG Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -|- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000229; CAB56375.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 354 AA; 39578 MW; 0B755E190255F33E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAAV 27
Db 339 LAALAAV 345
|||||

RESULT 536
Q82XV2 NITEU
ID Q82XV2 NITEU PRELIMINARY; PRT; 354 AA.
AC Q82XV2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=NE0147;
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
DOI=10.1128/JB.185.9.2759-2773.2003;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.L.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arciero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT *Complete genome sequence of the ammonia-oxidizing bacterium and
```

```
RT obligate chemolithoautotroph Nitrosomonas europaea.*;
RL J. Bacteriol. 185:2759-2773(2003).
DR EMBL; BX321856; CAD84058.1; -; Genomic_DNA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 354 AA; 37863 MW; AE6232C7EB134E90 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAAV 27
Db 180 LAALAAV 186
|||||

RESULT 537
Q7VE10 PROMA
ID Q7VE10 PROMA PRELIMINARY; PRT; 356 AA.
AC Q7VE10;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protein chain release factor B.
GN Name=prfB; OrderedLocNames=Pro0205;
OS Prochlorococcus marinus.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1219;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SARG / COMP 1375 / SS120;
RX MEDLINE=22810154; PubMed=12917486; DOI=10.1073/pnas.1733211100;
RA Dufrene A., Salanoubat M., Partensky F., Artiguenave F., Axmann I.M.,
RA Barbe V., Duprat S., Galperin M.Y., Koonin E.V., Le Gall F.,
RA Makarova K.S., Ostrowski M., Oztas S., Robert C., Rogozin I.B.,
RA Scanlan D.J., Tandeau de Marsac N., Weissenbach J., Wincker P.,
RA Wolf Y.I., Hess W.R.;
RT *Genome sequence of the cyanobacterium Prochlorococcus marinus SS120,
RT a nearly minimal oxyphototrophic genome.*;
RC Proc. Natl. Acad. Sci. U.S.A. 100:10020-10025(2003).
DR EMBL; AE017161; AAP99251.1; -; Genomic_DNA.
DR HSP; P07012; IQCE.
DR GO; GO:0005737; C:Cytoplasm; IEA.
DR GO; GO:0016149; F:translation release factor activity, codon . . .; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR GO; GO:0006415; P:translational termination; IEA.
DR InterPro; IPR005139; PCRF.
DR InterPro; IPR003352; PEP_rel_factor_I.
DR InterPro; IPR004374; PrfB.
DR InterPro; IPR012086; Release_factor.
DR Pfam; PF03462; PCRF; 1.
DR Pfam; PF00472; RF-1; 1.
DR PIRSF; PIRSF003056; Release_factor; 1.
DR TIGRFAMs; TIGR00020; prfB; 1.
DR PROSITE; PS00745; RF_PROK_I; 1.
KW Complete proteome.
SQ SEQUENCE 356 AA; 40040 MW; 1D68538C60D8ED8B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 DLEVTS 12
Db 219 DLEVTS 225
|||||

RESULT 538
Q92T67 RHIME
ID Q92T67 RHIME PRELIMINARY; PRT; 357 AA.
AC Q92T67;
```

DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE PUTATIVE OXIDOREDUCTASE PROTEIN.
 GN OrderedLocusNames=RO0113; ORFNames=SMC04139;
 OS Rhizobium meliloti (Sinorhizobium meliloti);
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
 RA Capela D., Barloy-Hubler F., Guzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puehler A., Fumelle B., Ramsparger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591782; CAC41500.1; -: Genomic DNA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR000683; GFO/IDH/MocA_N.
 DR InterPro; IPR004104; GFO/IDH/MocA_C.
 DR Pfam; PF01408; GFO_IDH_MocA_1.
 DR Pfam; PF02894; GFO_IDH_MocA_C; 1.
 KW Complete proteome.
 SQ SEQUENCE 357 AA; 39706 MW; 94A543A9036E40B5 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 357;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 20 VLAALAA 26
 Db 83 VLAALAA 89
 |||||

 RESULT 539
 ID Q7WH91 BORBR PRELIMINARY; PRT; 360 AA.
 AC Q7WH91;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BB3318;
 OS Bordetella bronchiseptica (Alcaligenes bronchiseptica).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cadeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640417; CAE33810.1; -: Genomic DNA.
 DR InterPro; IPR002549; UPP0118.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 361 AA; 38085 MW; A68FE1705B0808AFA CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 361;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 17 LGGVLA 23
 Db 325 LGGVLA 331
 |||||

 RESULT 541
 ID Q7WH91 BORPA PRELIMINARY; PRT; 361 AA.
 AC Q7WH91;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BP1789;
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

DR Pfam; PF01594; UPP0118; 1.
 KW Complete proteome.
 SQ SEQUENCE 360 AA; 37912 MW; 013B2405336062F1 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 360;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 17 LGGVLA 23
 Db 324 LGGVLA 330
 |||||

 RESULT 540
 ID Q7WH91 BORPE PRELIMINARY; PRT; 361 AA.
 AC Q7WH91;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BP2180;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi1227;
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cadeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.,
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640417; CAE342458.1; -: Genomic DNA.
 DR InterPro; IPR002549; UPP0118.
 DR Pfam; PF01594; UPP0118; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 361 AA; 38085 MW; A68FE1705B0808AFA CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 361;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 17 LGGVLA 23
 Db 325 LGGVLA 331
 |||||

 RESULT 541
 ID Q7WH91 BORPA PRELIMINARY; PRT; 361 AA.
 AC Q7WH91;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BP1789;
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achten M., Atkin R., Baker S., Basham D., Baoun N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders R., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640428; CA837091.1; -; Genomic_DNA.
DR InterPro; IPR002549; UPF0118; 1.
DR Pfam; PF01594; UPF0118; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 361 AA; 38043 MW; A585C3585122A2FA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGVLA 23
Db 325 LGVLA 331
|||||

RESULT 542
QAND64_9MICC PRELIMINARY; PRT; 362 AA.
AC QAND64;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Threonine aldolase (EC 4.1.2.5).
GN ORFNames=ArthDRAPT1069;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Micrococciaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Terani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHQ01000012; EAL95329.1; -; Genomic_DNA.
KW Lyase.
SQ SEQUENCE 362 AA; 39129 MW; 7D2BCE140B32EBE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 36 VLAALAA 42
|||||

```

```

RESULT 543
OS4158_STRCO PRELIMINARY; PRT; 362 AA.
AC OS4158;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative methyltransferase.
GN OrderedLocuNames=SC05895; ORFNames=SC3F7.15;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cardeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornaby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939125; CAAL16186.1; -; Genomic_DNA.
DR PIR; T34921.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR InterPro; IPR000780; Chet_mtfase.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR000051; SAM_bind.
DR SMART; SM00138; Metrc; 1.
KW Complete proteome; Methyltransferase; Transferase.
SQ SEQUENCE 362 AA; 39393 MW; ABE8FE4C0A3E3CF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 362;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 60 GVLAALA 66
|||||

RESULT 544
Q4HFY0_CAMCO PRELIMINARY; PRT; 364 AA.
AC Q4HFY0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Methyl-accepting chemotaxis transducer (TipC).
GN Name=tlpC-2; ORFNames=CC00539;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RM2228;
RA Pouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the

```


RT genomes of multiple Campylobacter species."; RT
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL: AAPL01000004; EAL56766.1; -; Genomic DNA.
 SQ SEQUENCE 364 AA; 40300 MW; 5DCAP08CBCCA25A2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
 Db 58 VLLGGVL 64

RESULT 545
 Y1821_SVNY3
 ID Y1821_SVNY3 STANDARD; PRT; 366 AA.
 AC P73714;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Hypothetical zinc metalloprotease slr1821 (EC 3.4.24.-).
 GN OrderedLocustNames=slr1821;
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=11148;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugitani M., Sasamoto S., Kimura T.,
 RA Hoshino T., Hikosawa M., Sugitani M., Nakazaki N., Naruo K., Okumura S.,
 RA Shimizu S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -!- COPACTOR: Zinc (Probable).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC -!- SIMILARITY: Belongs to the peptidase M50B family.
 CC -!- SIMILARITY: Contains 1 PDZ (DHR) domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL: BA000022; BAA17761.1; -; Genomic DNA.
 DR PIR: S77203; S77203.
 DR MEROPS: M50.004; -;
 DR InterPro: IPR001478; PDZ.
 DR InterPro: IPR004387; Pept_M50_Zn.
 DR InterPro: IPR006025; Pept_M_Zn_BS.
 DR InterPro: IPR008915; Peptidase_M50.
 DR Pfam: PF02163; Peptidase_M50; 1.
 DR SMART: SM00228; PDZ; 1.
 DR TIGRFAMs: TIGR00054; Pept_M50_Zn; 1.
 DR PROSITE: PS50106; PDZ; FALSE NEG.
 DR PROSITE: PS00142; ZINC_PROTEASE; FALSE NEG.
 KW Complete proteome; Hydrolase; Hypothetical protein; Inner membrane;
 KW Membrane; Metal-binding; Metalloprotease; Protease; Transmembrane;
 KW Zinc.
 FT TRANSMEM 95 115 Potential.
 FT TRANSMEM 293 313 Potential.
 FT TRANSMEM 325 345 Potential.
 FT DOMAIN 106 188 PDZ.
 FT ACT_SITE 21 21 Potential.

FT METAL 20 20 Zinc (catalytic) (potential).
 FT METAL 24 24 Zinc (catalytic) (potential).
 SQ SEQUENCE 366 AA; 38982 MW; CCB373F5089255D6 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 366;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 Db 3 VLAALAA 9

RESULT 546
 Q682F3_ARATH PRELIMINARY; PRT; 367 AA.
 ID Q682F3_ARATH PRELIMINARY;
 AC Q682F3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Shrunk seed protein (SSB1).
 GN Name=At2g45690;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Totoki Y., Seki M., Ishida J., Nakajima M., Enju A., Kamiya A.,
 RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
 RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
 RA Satou M., Toyoda T., Konagaya A., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.;
 RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs."
 RT Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AK175414; BAD43177.1; -; mRNA.
 SQ SEQUENCE 367 AA; 41614 MW; 69DB49A7DCD01BAE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QQAVIEP 102
 Db 209 QQAVIEP 215

RESULT 547
 Q8S8S1_ARATH PRELIMINARY; PRT; 367 AA.
 ID Q8S8S1_ARATH PRELIMINARY;
 AC Q8S8S1;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 23, Last annotation update)
 DE Expressed protein (Shrunken seed protein) (SSB1).
 GN Name=At2g45690;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.J.,
 RA Wu D., Maiti R., Roming C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.B., Bowman C.L., White O., Nietman W.C., Fraser C.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP NUCLEOTIDE SEQUENCE.
 RA Totoki Y., Seki M., Iehida J., Nakajima M., Enju A., Kamiya A.,
 RA Narusaka M., Shin-i T., Nakagawa M., Sakamoto N., Oishi K., Kohara Y.,
 RA Kobayashi M., Toyoda A., Sakaki Y., Sakurai T., Iida K., Akiyama K.,
 RA Satou M., Toyoda T., Konegaya A., Carninci P., Kawai J.,
 RA Hayashizaki Y., Shinozaki K.,
 RT "Large-scale analysis of RIKEN Arabidopsis full-length (RAFL) cDNAs.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC003680; AM14899.1; -; Genomic_DNA.
 DR EMBL; AK175396; BAD43159.1; -; mRNA.
 SQ SEQUENCE 367 AA; 41610 MW; 9667897388D55EB9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 QQAVIEP 102
 |||||
 Db 209 QQAVIEP 215

RESULT 548
 Q9XEG0 ARATH PRELIMINARY; PRT; 367 AA.
 AC Q9XEG0;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE Shrunk seed protein.
 GN Names=SR81;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99212146; PubMed=10195899; DOI=10.1126/science.284.5412.328;
 RA Lin Y., Sun L., Nguyen L.V., Rachubinski R.A., Goodman H.M.,
 RT "The Pex1p homolog SR81 and storage organelle formation in
 Arabidopsis seeds.";
 RL Science 284:328-330(1999).
 DR EMBL; AF085354; AAD30661.1; -; mRNA.
 SQ SEQUENCE 367 AA; 41624 MW; 35EE727140C485B8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 QQAVIEP 102
 |||||
 Db 209 QQAVIEP 215

RESULT 549
 Q9YCN9 AERPE PRELIMINARY; PRT; 370 AA.
 AC Q9YCN9;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein APE1219.
 GN OrderedLocusNames=APE1219;
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococcaeae; Aeropyrum.
 OC NCBI_TaxID=56636;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=KJ;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,

RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Maeda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 DR EMBL; BA000002; BAA80208.1; -; Genomic_DNA.
 DR PIR; B72594; B72594.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 370 AA; 38645 MW; 23A82D06C329F98E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 342 VLAALAA 348

RESULT 550
 Q7U3W6 SYNXP PRELIMINARY; PRT; 374 AA.
 AC Q7U3W6;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Peptide chain release factor RP-2.
 GN Name=prfB; OrderedLocusNames=SYNW2311;
 OS Synechococcus sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OC NCBI_TaxID=84588;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
 RA Palenik B., Brabham B., Larimer F.W., Land M.L., Hauser L.,
 RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,
 RA Paulsen I.T., Dufréne A., Partensky F., Webb E.A., Waterbury J.,
 RT "The genome of a motile marine Synechococcus.";
 RL Nature 424:1037-1042(2003).
 DR EMBL; BX569695; CA80826.1; -; Genomic_DNA.
 DR HSP; P07012; IQQE.
 DR GO; GO:0005737; Cytoplasm; IEA.
 DR GO; GO:0016149; F:translation release factor activity, codon . . .; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR GO; GO:0006415; P:translational termination; IEA.
 DR InterPro; IPR005139; PCRF.
 DR InterPro; IPR000352; PEP_rel_factor_I.
 DR InterPro; IPR004374; PrfB.
 DR Pfam; PF03462; PCRF; 1.
 DR Pfam; PF00472; RF-1; 1.
 DR PIRSF; PIRSF003056; Release_factor; 1.
 DR TIGRFAMs; TIGRF00020; prfB; 1.
 DR PROSITE; PS00745; RF_PROK_I; 1.
 KW Complete proteome.
 SQ SEQUENCE 374 AA; 41775 MW; A1PFC8CE06PBB444 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 374;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVTS 12
 |||||
 Db 237 DLEVTS 243

RESULT 551
 Q4LX47 9BURK PRELIMINARY; PRT; 375 AA.
 ID Q4LX47 9BURK PRELIMINARY;
 AC Q4LX47;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)


```

QY      17 LGGVLA 23
DB      288 LGGVLA 294
|||||
|||||

RESULT 554
Q6NAN4_RHOPA
ID      Q6NAN4_RHOPA PRELIMINARY; PRT; 382 AA.
AC      Q6NAN4;
DT      03-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT      05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE      Putative histidyl-tRNA synthetase precursor.
GN      OrderedLocusNames=RPAL150;
OS      Rhodopseudomonas palustris.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Rhodopseudomonas.
ON      NCBI_TaxID=1076;
RX      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CGA009 / ATCC BAA-98;
RX      PubMed=14704707; DOI=10.1038/nbr923;
RA      Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA      Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA      Gibson J.L., Hanson T.E., Bobet C., Torres y Torres J.L., Peres C.,
RA      Harrison P.H., Gibson J., Harwood C.S.;
RT      "Complete genome sequence of the metabolically versatile
RT      photosynthetic bacterium Rhodopseudomonas palustris."
RL      Nat. Biotechnol. 22:55-61(2004).
CC      -|- SUBUNIT: Homodimer (By similarity).
CC      -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -|- SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase
CC      family.
DR      EMBL; BX572596; CAB26593.1; -; Genomic_DNA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0016874; F:ligase activity; IEA.
DR      GO; GO:0004812; F:tRNA ligase activity; IEA.
DR      GO; GO:0006412; P:protein biosynthesis; IEA.
DR      GO; GO:0006418; P:tRNA aminoacylation for protein translation; IEA.
DR      InterPro; IPR003114; tRNA-synt_2b.
DR      InterPro; IPR006195; tRNA-synt_2b.
DR      Pfam; PF00587; tRNA-synt_2b; 1.
DR      PROSITE; PS00862; AA tRNA LIGASE II; 1.
DR      ATP-binding; Aminoacyl-tRNA synthetase; Complete proteome; Ligase;
DR      Nucleotide-binding; Protein biosynthesis; Signal.
FT      SIGNAL 1 17 Potential.
SQ      SEQUENCE 382 AA; 40285 MW; 62C636335070AFC0 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GVLAALA 25
DB      200 GVLAALA 206
|||||
|||||

RESULT 555
ID01739_CABEL
ID      001739_CABEL PRELIMINARY; PRT; 383 AA.
AC      001739;
DT      01-JUL-1997 (TrEMBLrel. 04, Created)
DT      01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein F20H11.5.
GN      ORFNames=F20H11.5;
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
ON      NCBI_TaxID=6239;
RX      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
SQ      SEQUENCE 383 AA; 41046 MW; A26EAC8A8F9BDBE1 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 382;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      19 GVLAALA 25
DB      200 GVLAALA 206
|||||
|||||

RESULT 555
ID01739_CABEL
ID      001739_CABEL PRELIMINARY; PRT; 383 AA.
AC      001739;
DT      01-JUL-1997 (TrEMBLrel. 04, Created)
DT      01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical protein F20H11.5.
GN      ORFNames=F20H11.5;
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Peloderinae; Caenorhabditis.
ON      NCBI_TaxID=6239;
RX      [1]
RP      NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
SQ      SEQUENCE 383 AA; 41046 MW; A26EAC8A8F9BDBE1 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 LLOGVLA 22
DB      209 LLOGVLA 215
|||||
|||||

RESULT 556
Q89DB5_BRAJA
ID      Q89DB5_BRAJA PRELIMINARY; PRT; 383 AA.
AC      Q89DB5;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hisz protein.
GN      Names=hisZ; OrderedLocusNames=blr7524;
OS      Bradyrhizobium japonicum.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Bradyrhizobium.
ON      NCBI_TaxID=375;
RX      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=USDA 110;
RX      MEDLINE=22484998; PubMed=12597275;
RA      Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA      Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA      Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA      Tabata S.;
RT      "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT      Bradyrhizobium japonicum USDA110."
RL      DNA Res. 9:189-197(2002).
CC      -|- FUNCTION: Required for the first step of histidine biosynthesis.
CC      May allow the feedback regulation of ATP phosphoribosyltransferase
CC      activity by histidine (By similarity).
CC      -|- PATHWAY: Amino-acid biosynthesis; L-histidine biosynthesis; L-
CC      histidine from PRPP: step 1.
CC      -|- SUBUNIT: Heteromultimer composed of hisG and hisZ subunits (By
CC      similarity).
CC      -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR      EMBL; BA000040; BAC52789.1; -; Genomic_DNA.
DR      HSPSP; P60906; 1HTT.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004812; F:tRNA ligase activity; IEA.
DR      GO; GO:0000105; P:histidine biosynthesis; IEA.
DR      GO; GO:0006418; P:tRNA aminoacylation for protein translation; IEA.
DR      InterPro; IPR003114; tRNA-synt_2b.
DR      Pfam; PF00587; tRNA-synt_2b; 1.
DR      Amino-acid biosynthesis; Complete proteome; Histidine biosynthesis.
KW      SEQUENCE 383 AA; 41046 MW; A26EAC8A8F9BDBE1 CRC64;

```

```

RC      STRAIN=Bristol N2;
RX      MEDLINE=99069613; PubMed=9851916;
RG      The C. elegans sequencing consortium;
RT      "Genome sequence of the nematode C. elegans: a platform for
RT      investigating biology."
RL      Science 282:2012-2018(1998).
DR      EMBL; AF002197; AAB53982.1; -; Genomic_DNA.
DR      PIR; F88486; F88486.
DR      HSPSP; P00371; 1AN9.
DR      Ensembl; F20H11.5; Caenorhabditis elegans.
DR      WormBase; WBGene00017648; F20H11.5.
DR      WormPep; F20H11.5; CE09514.
DR      GO; GO:0003884; F:D-amino-acid oxidase activity; IEA.
DR      GO; GO:0016491; P:oxidoreductase activity; IEA.
DR      GO; GO:0006118; P:electron transport; IEA.
DR      InterPro; IPR006181; DAO.
DR      InterPro; IPR006076; Fad_oxred.
DR      Pfam; PF01266; DAO; 1.
DR      PROSITE; PS00677; DAO; 1.
KW      Complete proteome; Hypothetical protein.
SQ      SEQUENCE 383 AA; 42502 MW; CECAB1SAEC1E93B5 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      16 LLOGVLA 22
DB      209 LLOGVLA 215
|||||
|||||

RESULT 556
Q89DB5_BRAJA
ID      Q89DB5_BRAJA PRELIMINARY; PRT; 383 AA.
AC      Q89DB5;
DT      01-JUN-2003 (TrEMBLrel. 24, Created)
DT      01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hisz protein.
GN      Names=hisZ; OrderedLocusNames=blr7524;
OS      Bradyrhizobium japonicum.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Bradyrhizobiaceae; Bradyrhizobium.
ON      NCBI_TaxID=375;
RX      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=USDA 110;
RX      MEDLINE=22484998; PubMed=12597275;
RA      Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA      Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA      Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA      Tabata S.;
RT      "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT      Bradyrhizobium japonicum USDA110."
RL      DNA Res. 9:189-197(2002).
CC      -|- FUNCTION: Required for the first step of histidine biosynthesis.
CC      May allow the feedback regulation of ATP phosphoribosyltransferase
CC      activity by histidine (By similarity).
CC      -|- PATHWAY: Amino-acid biosynthesis; L-histidine biosynthesis; L-
CC      histidine from PRPP: step 1.
CC      -|- SUBUNIT: Heteromultimer composed of hisG and hisZ subunits (By
CC      similarity).
CC      -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
DR      EMBL; BA000040; BAC52789.1; -; Genomic_DNA.
DR      HSPSP; P60906; 1HTT.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004812; F:tRNA ligase activity; IEA.
DR      GO; GO:0000105; P:histidine biosynthesis; IEA.
DR      GO; GO:0006418; P:tRNA aminoacylation for protein translation; IEA.
DR      InterPro; IPR003114; tRNA-synt_2b.
DR      Pfam; PF00587; tRNA-synt_2b; 1.
DR      Amino-acid biosynthesis; Complete proteome; Histidine biosynthesis.
KW      SEQUENCE 383 AA; 41046 MW; A26EAC8A8F9BDBE1 CRC64;

```

```

Query Match          5.9%; Score 7; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
DB 201 GVLAALA 207

RESULT 557
Q5FSL6 GLUOX PRELIMINARY; PRT; 384 AA.
ID Q5FSL6
AC Q5FSL6
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=GOX0856;
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Gluconobacter.
OX NCBI_TaxID=442;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=621H;
RX PubMed=15665824; DOI=10.1038/nbt1062;
RA Prust C., Hoffmeister M., Liesegang H., Wierze A., Fricke W.F.,
RA Ehrenreich A., Gottschalk G., Deppenmeier U.;
RT "Complete genome sequence of the acetic acid bacterium Gluconobacter
RT oxydans.";
RL Nat. Biotechnol. 23:195-200(2005).
DR EMBL; CP000009; AAW60630.1; -; Genomic_DNA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 384 AA; 41288 MW; B66CE83AF8CE7DA0 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 384;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 364 VLAALAA 370

RESULT 558
Y2056 BDEBA
ID Y2056 BDEBA STANDARD; PRT; 388 AA.
AC Q6MLF5;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical RNA methyltransferase Bg2056 (EC 2.1.1.-).
GN OrderedLocusNames=Bg2056;
OS Bdellovibrio bacteriovorus.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Bdellovibrionales;
OC Bdellovibrionaceae; Bdellovibrio.
OX NCBI_TaxID=959;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=HD100 / DSM 50701 / ATCC 15356 / NCIB 9529;
RX PubMed=14752164; DOI=10.1126/science.1093027;
RA Rendulic S., Jagtap P., Rosinus A., Eppinger M., Baar C., Lanz C.,
RA Keller H., Lambert C., Evans R.J., Goessmann A., Meyer F.,
RA Sockett R.E., Schuster S.C.;
RT "A predator unmasked: life cycle of Bdellovibrio bacteriovorus from a
RT genomic perspective.";
RL Science 303:689-692(2004).
CC -!- SIMILARITY: Belongs to the RNA M5U methyltransferase family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; BX842651; CAE79902.1; -; Genomic_DNA.
DR InterPro; IPR001737; RNA_meth_trans.
DR InterPro; IPR000051; SAM_Bd.
DR InterPro; IPR010280; U5_mtfase.
DR Pfam; PF05958; tRNA_U5-meth tr; 1.
DR PROSITE; PS01131; RENA_A_DIMETH; 1.
DR PROSITE; PS01230; TRMA_1; FALSE_NEG.
DR PROSITE; PS01231; TRMA_2; 1.
KW Complete proteome; Hypothetical protein; Methyltransferase;
FT ACT SITE 343 343 By similarity.
SQ SEQUENCE 388 AA; 44444 MW; B4343F240BB2CA9B CRC64;

Query Match          5.9%; Score 7; DB 1; Length 388;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VLGILQR 92
DB 379 VLGILQR 385

RESULT 559
O04178 BRACM
ID O04178 BRACM PRELIMINARY; PRT; 389 AA.
AC O04178;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Aminoalcoholphosphotransferase.
GN Name=AAPII;
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=L;
RA Min K.M., Cho S.H.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U96713; AAB53764.1; -; mRNA.
DR PIR; T14412; T14412.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000462; CDP-OH_P_trans.
DR Pfam; PF01066; CDP-OH_P_transf; 1.
DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
KW Transferase.
SQ SEQUENCE 389 AA; 43512 MW; F69C2EB6CBC363A CRC64;

Query Match          5.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
DB 264 VLLGGVL 270

RESULT 560
O82567 ARATH
ID O82567 ARATH PRELIMINARY; PRT; 389 AA.
AC O82567;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)

```

DE Aminoalcoholphosphotransferase.
GN Name=AAPT1; Synonyms=Atlg133560, F13B4.5;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Dewey R.B., Goode J.H.;
RT "Characterization of aminoalcoholphosphotransferases from Arabidopsis
thaliana and soybean";
RL Plant Physiol. Biochem. 37:445-457 (2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carmichael P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.P.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler B., Chao Q., Chin C., Chiu J., Choi E., Gonzalez A.,
RA Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharly N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vayenberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carmichael P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA EMBL; AF091843; AAC61768.1; -; mRNA.
DR EMBL; AY080725; AAL86327.1; -; mRNA.
DR EMBL; AC027134; AAF99823.1; -; Genomic_DNA.
DR EMBL; AY114062; AAM45110.1; -; mRNA.
DR PIR; F86268; F86268.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000463; CDP-OH_P_trans.
DR Pfam; PF01066; CDP-OH_P_transf; 1.
DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
KW Transferase.
SQ SEQUENCE 389 AA; 43542 MW; 2B91B905E14B9D17 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 VLLGGVLL 21
DB 264 VLLGGVLL 270
RESULT 561
Q9SP57_BRACM
ID Q9SP57_BRACM PRELIMINARY; PRT; 389 AA.
AC Q9SP57;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Aminoalcoholphosphotransferase.
OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=3711;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA MEDLINE=20550483; PubMed=11100782; DOI=10.1093/pcp/pcd030;
RA Choi Y.H., Lee J.K., Lee C.H., Cho S.H.;
RT "cDNA cloning and expression of an aminoalcoholphosphotransferase
isoform in Chinese cabbage";
RL Plant Cell Physiol. 41:1080-1084 (2000).
DR EMBL; AF183933; AAD56040.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0008654; P:phospholipid biosynthesis; IEA.
DR InterPro; IPR000463; CDP-OH_P_trans.
DR Pfam; PF01066; CDP-OH_P_transf; 1.
DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
KW Transferase.
SQ SEQUENCE 389 AA; 43259 MW; 0F1E86E440CDA6B3 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 15 VLLGGVLL 21
DB 264 VLLGGVLL 270
RESULT 562
Q912P7_PSEAB
ID Q912P7_PSEAB PRELIMINARY; PRT; 389 AA.
AC Q912P7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Probable MFS transporter.
GN OrderedLocustNames=PA1848;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RA MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warriner P.,
RA Hickey M.J., Brinkman F.S.L., Huftnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H. Jr., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen";
RL Nature 406:959-964 (2000).
DR EMBL; AE004611; AAG05237.1; -; Genomic_DNA.
DR PIR; G83413; G83413.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS 1.
DR Pfam; PF07690; MFS 1; 1.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 389 AA; 40754 MW; BB90A908F5A1DE51 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 389;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24
Db 140 GGVLAAL 146

RESULT 563
Q9YDB2 AERP8
ID Q9YDB2 AERP8 PRELIMINARY; PRT; 390 AA.
AC Q9YDB2;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DE Hypothetical protein APL1001.
GN OrderedlocusNames=APE1001;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococcales; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Aikai A., Kosugi H.,
RA Hosoyma A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Rep. 6:83-101(1999).
DR EMBL; BA000002; BAA79985.1; -; Genomic_DNA.
DR PIR; A72698; A72698.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 390 AA; 41090 MW; 9E98D66EEAEFD207 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 390;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 266 VLAALAA 272

RESULT 564
Q51RJ1 MAGGR
ID Q51RJ1 MAGGR PRELIMINARY; PRT; 392 AA.
AC Q51RJ1;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Predicted protein.
GN ORFNames=MG09760.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaportha.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambuster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Biye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collimore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,

RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kamat M., Kamat A., Kamysseles M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokytang T., Lokytang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Maucelli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihaliev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., News C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okaowo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schubach R., Seaman C., Settillalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,
RA Towey S., Tsamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea";
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACU01001298; EAA49769.1; -; Genomic_DNA.
SQ SEQUENCE 392 AA; 43084 MW; 1BD22F7423345C4C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 104 VLAALAA 110

RESULT 565
Q8W1E9 BRARP
ID Q8W1E9 BRARP PRELIMINARY; PRT; 392 AA.
AC Q8W1E9;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Aminoalcoholphosphotransferase.
GN Name=AAPT3;
OS Brassica rapa (subsp. pekinensis) (Chinese cabbage).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_TaxID=51351;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kim K.S., Cho S.H.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF446089; AAL46934.3; -; mRNA.

```
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016740; P: transferase activity; IEA.
DR GO; GO:0008654; P: phospholipid biosynthesis; IEA.
DR InterPro; IPR000462; CDP-OH_P_trans.
DR Pfam; PF01066; CDP-OH_P_trans; 1.
DR PROSITE; PS00379; CDP_ALCOHOL_P_TRANSF; 1.
KW Transferase.
SQ SEQUENCE 392 AA; 43827 MW; ALEDECE3437D8798F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVLL 21
DB 264 VLLGGVLL 270

RESULT 566
Q5YVF7_NOCFA PRELIMINARY; PRT; 392 AA.
AC Q5YVF7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative transporter.
GN OrderedLocuNames=nfa29870;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD57834.1; -; Genomic_DNA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0005215; P: transporter activity; IEA.
DR GO; GO:0046677; P: response to antibiotic; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 392 AA; 38848 MW; 6DB2485C88D0E307 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 392;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 99 VLAALAA 105

RESULT 567
Q5AC31_CANAL PRELIMINARY; PRT; 394 AA.
AC Q5AC31;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein GCV1.
GN Names=GCV1; ORFNames=CaO19.5519;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
```

```
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans."
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACQ01000034; EAL00186.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 394 AA; 43639 MW; 314322EC43959282 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 HIELGGK 45
DB 25 HIELGGK 31

RESULT 568
Q5ACF2_CANAL PRELIMINARY; PRT; 394 AA.
AC Q5ACF2;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein GCV1.
GN Names=GCV1; ORFNames=CaO19.12965;
OS Candida albicans SC5314.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=237561;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RX PubMed=15123810; DOI=10.1073/pnas.0401648101;
RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,
RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,
RA Davis R.W., Scherer S.;
RT "The diploid genome sequence of Candida albicans."
RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SC5314;
RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,
RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,
RA Jones T., Scherer S., Agabian N.;
RT "Annotation of the Genome of Candida albicans."
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACQ01000033; EAL00308.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 394 AA; 43768 MW; 707612340651D9EB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
QY 39 HIELGGK 45
Db 25 HIELGGK 31

RESULT 569
Q5H429_XANOR PRELIMINARY; PRT; 394 AA.
AC Q5H429;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Murein hydrolase D.
GN Name=dniR; OrderedLocusNames=XO01038;
OS Xanthomonas oryzae (pv. oryzae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=64187;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=KACC10331 / KX085;
RX PubMed=15673718; DOI=10.1093/nar/gki206;
RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
RA Go S.-J.;
RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
RT the bacterial blight pathogen of rice.";
RL Nucleic Acids Res. 33:577-586(2005).
DR EMBL; AB013598; AAW4292.1; -; Genomic DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR InterPro; IPR008258; LT catalytic.
DR InterPro; IPR002482; LyxM.
DR InterPro; IPR002155; Thiolase.
DR Pfam; PF01476; LyxM; 1.
DR Pfam; PF01464; SLT; 1.
DR SMART; SM00257; LyxM; 1.
DR PROSITE; PS00099; THIOLASE_3; UNKNOWN_1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 394 AA; 41530 MW; D91023BB43F5EA0E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 394;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 7 VLAALAA 13

RESULT 570
Q579F3_BRUAB PRELIMINARY; PRT; 396 AA.
AC Q579F3;
DT 10-MAY-2005 (TREMBlrel. 30, Created)
DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
DE Sugar ABC transporter, permease protein.
GN OrderedLocusNames=Brub2_0298;
OS Brucella abortus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=235;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=9-941 / Biovar 1;
RX PubMed=15805518; DOI=10.1128/JB.187.8.2715-2726.2005;
RA Halling S.M., Peterson-Burch B.D., Bricker B.J., Zuerner R.L.,
RA Qing Z., Li L.-L., Kapur V., Alt D.P., Olsen S.C.;
RT "Completion of the genome sequence of Brucella abortus and comparison
RT to the highly similar genomes of Brucella melitensis and Brucella

suis.";
RL J. Bacteriol. 187:2715-2726(2005).
DR EMBL; AB017224; AAX75731.1; -; Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 396 AA; 41970 MW; 80E9847D1085EC0B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
Db 301 GVLAALA 307

RESULT 571
Q9RS23_DEIRA PRELIMINARY; PRT; 396 AA.
AC Q9RS23;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein DR2304.
GN OrderedLocusNames=DR2304;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey B.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Morfat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AB002062; AAF11854.1; -; Genomic_DNA.
DR PIR; B75290; B75290.
DR TIGR; DR2304; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 396 AA; 39229 MW; 355FCB5E0179F312 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLA 22
Db 160 LGGVLA 166

RESULT 572
Q89QB8_BRAJA PRELIMINARY; PRT; 396 AA.
AC Q89QB8;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Sugar ABC transporter permease protein.
GN OrderedLocusNames=blr3210;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
```

```

OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriiguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpou S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC48475.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 1.
KW Complete proteome.
SQ SEQUENCE 396 AA; 41743 MW; A7213C7974E2F135 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 301 GVLAALA 307
|||||

RESULT 574
Q8YD15_BRUME
ID Q8YD15_BRUME PRELIMINARY; PRT; 396 AA.
AC Q8YD15;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE XLOSE TRANSPORT SYSTEM PERMEASE PROTEIN XYLH.
DR OrderedLocusNames=BMEI10362;
GN Brucella melitensis.
OS Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muler C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009673; AAL53604.1; -; Genomic_DNA.
DR PIR; A13554; A13554.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR InterPro; IPR001851; Bac_inmem_transp.
DR Pfam; PF02653; BPD_transp_2; 1.
KW Complete proteome.
SQ SEQUENCE 396 AA; 41990 MW; 586AB7564CB129D0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 301 GVLAALA 307
|||||

RESULT 575
Q5YU07_NOCFA
ID Q5YU07_NOCFA PRELIMINARY; PRT; 396 AA.
AC Q5YU07;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative two-component system sensor kinase.
GN OrderedLocusNames=hfal7980;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.

```

OX NCBI_TaxID=37329;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of *Nocardia farcinica* IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
CC -!- FUNCTION: Member of the two-component regulatory system lytR/lytS
CC that regulates genes involved in autolysis and cell wall
CC metabolism. Regulates the activity of the cell wall-associated
CC murein hydrolase through regulation of lrgA and lrgB (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC EMBL; AP006618; BAD56644.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016301; F:kinase activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0016310; F:phosphorylation; IEA.
DR GO; GO:0000160; F:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; ATP bd ATPase.
DR InterPro; IPR010559; His_kin_int.
DR InterPro; IPR004358; His_kin_like_C.
DR Pfam; PF02518; HATPase_c; 1.
DR Pfam; PF06580; His_kinase; 1.
DR PRINTS; PR00344; BCTRLSENSOR.
DR SMART; SM00387; HATPase_c; 1.
KW Complete proteome; Kinase; Transferase; Transmembrane;
KW Two-component regulatory system; B737F74BAF56844F CRC64;
SQ SEQUENCE 396 AA; 41272 MW; B737F74BAF56844F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 10 VLAALAA 16

RESULT 576
Q7WHY1 BORR
ID Q7WHY1 BORR PRELIMINARY; PRT; 398 AA.
AC Q7WHY1;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE Hypothetical protein.
GN OrderedLocustNames=BB3125;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Holden M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Helden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham K., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagsis K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares K.,
RA Unwin L., Whitehead S., Barrell B.G., Maekell D.J.;
RT "Comparative analysis of the genome sequences of *Bordetella pertussis*,
RT *Bordetella parapertussis* and *Bordetella bronchiseptica*.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640446; CAE33617.1; -; Genomic DNA.

DR HSP; P77407; IQ7E.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR003673; CAIB BAIF.
RX Pfam; PF02515; CoA_transf_3; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 398 AA; 42350 MW; C18FDEA013A85E3A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 398;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
| | | | |
Db 181 GVLAALA 187

RESULT 577
Q8BXV8 MOUSE
ID Q8BXV8_MOUSE PRELIMINARY; PRT; 399 AA.
AC Q8BXV8;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:A730065C21 product:Ellis van Creveld gene
DE homolog (human), full insert sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wuzhuh-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

MDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayateu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RN Genome Res. 10:1617-1630(2000).
[5]
RN NUCLEOTIDE SEQUENCE
RP STRAIN=CS7BL/6J; TISSUE=Cerebellum;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Iwawa M., Ohara E., Watanabe M.,
RA Itoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:1757-1771(2000).
[6]
RN NUCLEOTIDE SEQUENCE
RC STRAIN=CS7BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayateu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kigawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki R., Murata M., Nakamura M.,
RA Nishii K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tegami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR ENMBL; AK043184; BAC31486.1; -, mRNA.
DR Ensembl; ENSMUSG0000029122; Mus musculus.
SQ SEQUENCE 399 AA; 44037 MW; 3065516B23DB90C18 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
| | | | |
Db 27 VLLGGVL 33

RESULT 578
Q6AU99 ORYSA
ID Q6AU99_ORYSA PRELIMINARY; PRT; 401 AA.
AC Q6AU99;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein P0668F02.2.
GN Name=P0668F02.2;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN NUCLEOTIDE SEQUENCE.
RP Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Heu C.-H., Huang J.-J., Kuo P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-P.;
RT "Oryza sativa PAC P0668F02 genomic sequence.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR ENMBL; AC130729; AAT85248.1; -, Genomic DNA.

Gramene; O6AU99; -;
DR InterPro; IPR002885; PPR.
DR Pfam; PF01535; PPR; 8.
DR TIGRFAMs; TIGR00756; PPR; 5.
KW Hypothetical protein.
SQ SEQUENCE 401 AA; 42988 MW; D2C33537CP8P8P61 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
| | | | |
Db 318 LGGVLA 324

RESULT 579
Q8KYU7_9PROT
ID Q8KYU7_9PROT PRELIMINARY; PRT; 403 AA.
AC Q8KYU7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Membrane protein, putative.
GN Name=EBAC000-60D04_94;
OS Uncultured proteobacterium.
OC Bacteria; Proteobacteria; environmental samples.
OX NCBI_TaxID=153809;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=21822632; PubMed=11832943; DOI=10.1038/415630a;
RA Beja O., Suzuki M.T., Heidelberg J.F., Nelson W.C., Preston C.M.,
RA Hamada T., Eisen J.A., Fraser C.M., DeLong E.F.;
RT "Unsuspected diversity among marine aerobic anoxygenic phototrophs.";
RL Nature 415:630-633(2002).
DR ENMBL; AE008921; AAM48737.1; -, Genomic DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transporter activity; IEA.
DR InterPro; IPR011701; MFS.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS00850; MFS; 1.
SQ SEQUENCE 403 AA; 43450 MW; E733F1F77367901C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 403;
Best Local Similarity 100.0%; Pred. No. 5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 82 VLAALAA 88

RESULT 580
O69757_PSEFL
ID O69757_PSEFL PRELIMINARY; PRT; 405 AA.
AC O69757;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical metabolite transport protein.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=294;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=23F;
RC MEDLINE=97444287; PubMed=9300819; DOI=10.1016/S0378-1119(97)00151-0;
RA Kulakova A.N., Kulakov L.A., Quinn J.P.;
RT "Cloning of the phosphonoacetate hydrolase gene from Pseudomonas

fluorescens 23F encoding a new type of carbon-phosphorus bond cleaving enzyme and its expression in *Escherichia coli* and *Pseudomonas putida*.";
 RT Gene 195:49-53 (1997).
 RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=23F;
 RX MEDLINE=21242724; PubMed=11344133;
 RX DOI=10.1128/JB.183.11.3268-3275.2001;
 RA Kulakova A.N., Kulakov L.A., Akulenko N.V., Ksenzenko V.N.,
 RA Hamilton J.T.G., Quinn J.P.;
 RT "Structural and functional analysis of the phosphonoacetate hydrolase (phnA) gene region in *Pseudomonas fluorescens* 23F.";
 RL J. Bacteriol. 183:3268-3275 (2001).
 DR EMBL; L49465; AAC15508.1; -; Genomic DNA.
 DR GO; GO:0019866; C:inner membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 405 AA; 41684 MW; 59CF90DC856B327D CRC64;
 [1]
 Query Match 5.9%; Score 7; DB 2; Length 405;
 Best Local Similarity 100.0%; Pred. No. 5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 GVLAALA 25
 Db 288 GVLAALA 294
 [1]
 RESULT 581
 Q8XH17 CLOPE
 ID Q8XH17 CLOPE PRELIMINARY; PRT; 408 AA.
 AC Q8XH17;
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Probable nucleoside transporter.
 GN OrderedLocusNames=CPE2496;
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1502;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=13 / Type A;
 RX MEDLINE=21664373; PubMed=11792842; DOI=10.1073/pnas.022493799;
 RX Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:996-1001 (2002).
 RL EMBL; BA000016; BAB82202.1; -; Genomic DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0001882; F:nucleoside binding; IEA.
 DR GO; GO:0005415; P:nucleoside:sodium symporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR008276; C_nucleid_transp.
 DR InterPro; IPR011642; Gate.
 DR InterPro; IPR011657; Nucleos_tra2 C.
 DR InterPro; IPR002668; Nucleid_transp2.
 DR Pfam; PF07670; Gate; 1.
 DR Pfam; PF07662; Nucleos_tra2 C; 1.
 DR Pfam; PF01773; Nucleos_tra2_N; 1.
 DR ProDom; PD003768; Nucleid_transp2; 1.

DR TIGRFAMs; TIGR00804; nup; 1.
 KW Complete proteome.
 SQ SEQUENCE 408 AA; 42442 MW; 2F23604EB4119210 CRC64;
 [2]
 Query Match 5.9%; Score 7; DB 2; Length 408;
 Best Local Similarity 100.0%; Pred. No. 5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 16 LLGGVLA 22
 Db 388 LLGGVLA 394
 [1]
 RESULT 582
 Q9HSM8 HALSA
 ID Q9HSM8 HALSA PRELIMINARY; PRT; 410 AA.
 AC Q9HSM8;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Oxalate/formate antiporter.
 GN Name=oxlt; OrderedLocusNames=VNG0157G;
 OS Halobacterium salinarum (Halobacterium halobium).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=2242;
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
 RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
 NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Laskey S.R., Balliga N.S., Thorsen V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of *Halobacterium* species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181 (2000).
 DR EMBL; AB004982; AAG18775.1; -; Genomic DNA.
 DR PIR; C84176; C84176.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 KW Complete proteome.
 SQ SEQUENCE 410 AA; 42467 MW; 5E7EADFD99AFB50 CRC64;
 [1]
 Query Match 5.9%; Score 7; DB 2; Length 410;
 Best Local Similarity 100.0%; Pred. No. 5.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 16 LLGGVLA 22
 Db 389 LLGGVLA 395
 [1]
 RESULT 583
 Q65CE8 9ACTO
 ID Q65CE8 9ACTO PRELIMINARY; PRT; 413 AA.
 AC Q65CE8;
 DT 25-OCT-2004 (TRENBLrel. 28, Created)
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
 DE Putative sugar ABC-transporter permease.
 GN Name=kanQ;
 OS Streptomyces kanamyceticus.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.

```

OX NCBI_TaxID=1967;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC12853;
RX PubMed=15313224; DOI=10.1016/j.abb.2004.06.009;
RL Kharel M.K., Subba B., Basnet D.B., Wood J.S., Lee H.C., Liou K.,
RA Sohng J.K.;
RT "A gene cluster for biosynthesis of kanamycin from Streptomyces
RT kanamyceticus: comparison with gentamicin biosynthetic gene cluster.";
RL Arch. Biochem. Biophys. 429:204-214 (2004).
DR EMBL; AJ582817; CAF60514.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001851; SBC_inmem_transp.
DR InterPro; IPR000914; SBP_dac_5.
DR Pfam; PF02653; BPD_transp_1.
DR DR PROSITE; PS01040; SBP_BACTERIAL_5; UNKNOWN_1.
SQ SEQUENCE 413 AA; 43426 MW; 0ADA528A55ECB45 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 315 GVLAALA 321
|||||

RESULT 584
Q88HK7_PSEPK PRELIMINARY; PRT; 413 AA.
AC Q88HK7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Major facilitator family transporter.
DE OrderedLocustNames=PP3349;
GN Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
RX DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Daugherty S.C., Kolonay J.F.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Khouri H.M.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Tran K.,
RA Hance I., Chris Lee P., Holtzapfle E.K., Scanlan D., Tran K.,
RA Moazzaz A., Utterback T.R., Rizzo M., Lee K., Kossack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AE016786; AAN68953.1; -; Genomic_DNA.
DR TIGR; PP3349; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS0850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 41918 MW; CF9DC8138D144A59 CRC64;

```

```

Query Match 5.9%; Score 7; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 383 VLAALAA 389
|||||

RESULT 585
Q7NZQ6_CHRVO PRELIMINARY; PRT; 413 AA.
AC Q7NZQ6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable dipeptidyl aminopeptidase.
DE OrderedLocustNames=CV0865;
GN Chromobacterium violaceum.
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Chromobacterium.
OX NCBI_TaxID=536;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ATCC 12472 / DSM 30191;
RX MEDLINE=22882880; PubMed=14500782; DOI=10.1073/pnas.1832124100;
RA Vasconcelos A.T.R., de Almeida D.F., Hungria M., Guimaraes C.T.,
RA Antonio R.V., Almeida F.C., de Almeida L.G.P., de Almeida R.,
RA Alves-Gomes J.A., Andrade E.M., Araripe J., de Araujo M.F.F.,
RA Astolfi-Filho S., Azevedo V., Baptista A.J., Bataus L.A.M.,
RA Batista J.S., Beio A., van den Berg C., Bogo M., Bonatto S.,
RA Bordignon J., Brigido M.M., Brito C.A., Brocchi M., Burity H.A.,
RA Camargo A.A., Cardoso D.D.P., Carneiro N.P., Carraro D.M.,
RA Carvalho C.M.B., Cascado J.C.M., Cavada B.S., Chueire L.M.O.,
RA Creczynski-Pasa T.B., Cunha-Junior N.C., Fagundes N., Falcao C.L.,
RA Fantinatti F., Farias I.P., Felipe M.S.S., Ferrari L.P., Ferro J.A.,
RA Ferro M.I.T., Franco G.R., Freitas N.S.A., Furian L.N.,
RA Gazzinelli R.T., Gomes E.A., Goncalves P.R., Grangeiro T.B.,
RA Grattapaglia D., Grissard E.C., Hanna E.S., Jardim S.N., Laurino J.,
RA Leoi L.C.T., Lima L.F.A., Loureiro M.F., Lyra M.C.C.P.,
RA Madeira H.M.F., Manfio G.P., Maranhao A.O., Martins W.S.,
RA di Mauro S.M.Z., de Medeiros S.R.B., Meissner R.V., Moreira M.A.M.,
RA Nascimento F.P., Nicolas M.F., Oliveira J.G., Oliveira S.C.,
RA Paixao R.F.C., Parente J.A., Pedrosa F.O., Pena S.D.J., Pereira J.O.,
RA Pereira M., Pinto L.S.R.C., Pinto L.S., Porto J.I.R., Potrich D.P.,
RA Ramalho-Neto C.E., Reis A.M.M., Rigo L.U., Rondinelli E.,
RA Santos E.B.P., Santos F.R., Schneider M.P.C., Seunanez H.N.,
RA Silva A.M.R., da Silva A.L.C., Silva D.W., Silva R., Simoes I.C.,
RA Simon D., Soares C.M.A., Soares R.B.A., Souza E.M., Souza K.R.L.,
RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
RA Vetore A., Wassem R., Zaha A., Simpson A.J.G.;
RT "The complete genome sequence of Chromobacterium violaceum reveals
RT remarkable and exploitable bacterial adaptability.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665 (2003).
DR EMBL; AE016912; AAQ58540.1; -; Genomic DNA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR008391; AXE1.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF05448; AXE1; 1.
DR KW Aminopeptidase; Complete proteome; Hydrolase.
SQ SEQUENCE 413 AA; 45899 MW; 0F16E923978FB9C5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 413;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LALAAY 27
Db 253 LALAAY 259
|||||

```

```

RESULT 586
Q8QWV4 VIBCV
ID Q8QWV4 VIBCV PRELIMINARY; PRT; 414 AA.
AC Q8QWV4;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Multidrug resistance protein D.
GN OrderedLocusNames=VCA0214;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
NUCLEOTIDE SEQUENCE.
RP STRAIN=El Tor N16961 / Serotype O1;
RC MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.P., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004361; AAF96126.1; -; Genomic_DNA.
DR PIR; D82487; D82487.
DR TIGR; VCA0214; -.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS0850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 414 AA; 44106 MW; 3CB3C9AEB9C26A75 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 414;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 247 VLLGGVL 253

RESULT 587
Q8DA59 VIBVU
ID Q8DA59 VIBVU PRELIMINARY; PRT; 415 AA.
AC Q8DA59;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Multidrug resistance protein D.
GN OrderedLocusNames=VV12351;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
NUCLEOTIDE SEQUENCE.
RP STRAIN=CMCP6;
RC Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.B.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6."
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016804; AA010725.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.

Query Match 5.9%; Score 7; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 247 VLLGGVL 253

RESULT 588
Q7MK22 VIBVY
ID Q7MK22 VIBVY PRELIMINARY; PRT; 415 AA.
AC Q7MK22;
DT 01-MAR-2004 (TREMELrel. 26, Created)
DT 01-MAR-2004 (TREMELrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Permease.
GN OrderedLocusNames=VV1988;
OS Vibrio vulnificus (strain VJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
NUCLEOTIDE SEQUENCE.
RP PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C., Su T.-L.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
pathogen."
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC94752.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004812; Eflux_Bcr_CflA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR TIGRfams; TIGR00710; eflux_Bcr_CflA; 1.
DR PROSITE; PS0850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 415 AA; 43888 MW; FCBC77018D807A64 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 248 VLLGGVL 254

RESULT 589
Q82136 STRAW
ID Q82136 STRAW PRELIMINARY; PRT; 416 AA.
AC Q82136;
DT 01-JUN-2003 (TREMELrel. 24, Created)
```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DT Putative lipoprotein.
GN OrderedLocuNames=SAV3322;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=33903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
DR EMBL; BA000030; BAC71033.1; -; Genomic_DNA.
DR InterPro; IPR005490; ErfK_Ybis_YhnG.
DR Pfam; PF03734; ErfK_Ybis_YhnG; 1.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 416 AA; 43311 MW; D2E098441DB37816 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 15 VLLGGVL 21

RESULT 590
Q34184_CEPNE PRELIMINARY; PRT; 417 AA.
AC Q34184;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
OS Cepaea nemoralis (Banded wood snail).
OC Mitochondrion.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Helicoidea; Helicidae; Cepaea.
OC NCBI_TaxID=28835;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hepatopancreas;
RX MEDLINE=97077301; PubMed=8919868;
RA Terrett J.A., Miles S., Thomas R.H.;
RA "Complete DNA sequence of the mitochondrial genome of Cepaea nemoralis
RA (Gastropoda; Pulmonata).";
RL J. Mol. Evol. 42:160-168(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Hepatopancreas;
RX MEDLINE=97207849; PubMed=9055084;
RA Yamazaki N., Ueshima R., Terrett J.A., Yokobori S., Kaifu M.,
RA Segawa K., Kobayashi T., Numachi K., Ueda T., Nishikawa K.,
RA Watanabe K., Thomas R.H.;
RT "Evolution of pulmonate gastropod mitochondrial genomes: comparisons
RT of complete gene organizations of Euhadra, Cepaea and Albinaria and

```

```

RT implications of unusual tRNA secondary structures.";
RL Genetics 145:749-758(1997).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
DR EMBL; U23045; AAC09525.1; -; Genomic_DNA.
DR PIR; T11387; T11387.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0008137; P:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR InterPro; IPR003918; NADHub_oxred4.
DR InterPro; IPR001750; Oxidored_g1.
DR Pfam; PF00361; Oxidored_g1; 1.
DR PRINTS; PR01437; NUXDRDTASE4.
KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.
FT NON_TER 417 417
SQ SEQUENCE 417 AA; 46452 MW; B5B401DACD6E96F8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 235 GGVLAAL 241

RESULT 591
MURA ACIGB STANDARD; PRT; 419 AA.
ID MURA ACIGB STANDARD; PRT; 419 AA.
AC P33986;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE UDP-N-acetylglucosamine 1-carboxyvinyltransferase (EC 2.5.1.7)
DE (Enolpyruvate transferase) (UDP-N-acetylglucosamine enolpyruvyl
DE transferase) (EPT).
GN Name=murA; Synonyms=murZ;
OS Acinetobacter genomosp. 11.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OC NCBI_TaxID=106649;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RC STRAIN=ATCC 11171 / NCIB 8250 / CIP 63.46 / B94;
RX MEDLINE=94237446; PubMed=8181716; DOI=10.1016/0378-1097(94)90185-6;
RA Ehrt S., Hillen W.;
RT "UDP-N-acetylglucosamine 1-carboxyvinyl-transferase from Acinetobacter
RT calcoaceticus.";
RL FEMS Microbiol. Lett. 117:137-142(1994).
CC -1- FUNCTION: Cell wall formation. Adds enolpyruvyl to UDP-N-
CC acetylglucosamine. Target for the antibiotic phosphomycin.
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + UDP-N-acetyl-D-
CC glucosamine = phosphate + UDP-N-acetyl-3-O-(1-carboxyvinyl)-D-
CC glucosamine.
CC -1- PATHWAY: Peptidoglycan biosynthesis; first step.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: Belongs to the EPSP synthase family. MurA subfamily.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; L26051; AAA21618.1; -; Genomic_DNA.
CC HSPSP; P33038; 1DLG.
DR HAMAP; MF 00111; -; 1.
DR InterPro; IPR005750; AcGlu_Tran_MurA.
DR InterPro; IPR001986; EPSP_synth.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synth; 1.
DR TIGRFAMs; TIGR01072; murA; 1.
KW Cell cycle; Cell division; Cell shape; Cell wall;

```



```

KW Peptidoglycan synthesis; Transferase.
FT ACT_SITE 116 Proton donor (By similarity).
FT BINDING 116 PEP (covalent) (By similarity).
SQ SEQUENCE 419 AA; 44680 MW; E344CC2886C1BBE9 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 5.1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 378 VLAALAA 384

RESULT 592
Y1258 MYCTU STANDARD; PRT; 419 AA.
AC P64783; Q11060;
DT 01-OCT-1996 (Rel. 34, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DE Hypothetical protein Rv1258c/MT1297.
GN OrderedLocusNames=Rv1258c, MT1297; ORFNames=MTVCY50.24;
OS Mycobacterium tuberculosis
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekaita F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feldwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jegels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Uitterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.N.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL J. Bacteriol. 184:5479-5490 (2002).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX842576; CAB00894.1; -; Genomic DNA.
CC EMBL; AE000516; AAK45555.1; -; Genomic DNA.
CC PIR; B70753; B70753.
CC TIGR; MT1297; -.
CC Tuberculliet; Rv1258c; -.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR011701; MFS_1.
CC Pfam; PF07690; MFS_1; 1.

DR PROSITE; PS08050; MFS; 1.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 10 Potential.
FT TRANSMEM 46 66 Potential.
FT TRANSMEM 74 94 Potential.
FT TRANSMEM 103 123 Potential.
FT TRANSMEM 152 172 Potential.
FT TRANSMEM 173 193 Potential.
FT TRANSMEM 228 248 Potential.
FT TRANSMEM 263 283 Potential.
FT TRANSMEM 304 324 Potential.
FT TRANSMEM 348 368 Potential.
FT TRANSMEM 375 395 Potential.
SQ SEQUENCE 419 AA; 43287 MW; 69EA46BBFF653037 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 5.1e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 108 VLAALAA 114

RESULT 593
Y1288 MYCBO STANDARD; PRT; 419 AA.
AC P64784; Q11060;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE Hypothetical protein Mb1288c.
GN OrderedLocusNames=Mb1288c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis."
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882 (2003).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX248338; CAD94149.1; -; Genomic DNA.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR011701; MFS_1.
CC Pfam; PF07690; MFS_1; 1.
KW Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSMEM 10 30 Potential.
FT TRANSMEM 46 66 Potential.
FT TRANSMEM 74 94 Potential.
FT TRANSMEM 103 123 Potential.
FT TRANSMEM 152 172 Potential.
FT TRANSMEM 173 193 Potential.
FT TRANSMEM 228 248 Potential.
FT TRANSMEM 263 283 Potential.
FT TRANSMEM 304 324 Potential.
FT TRANSMEM 348 368 Potential.
FT TRANSMEM 375 395 Potential.

```

```

SQ SEQUENCE 419 AA; 43287 MW; 69EA46BBFF653037 CRC64;
Query Match 5.9%; Score 7; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 108 VLAALAA 114
|||||

RESULT 594
Q5LX88_SILPO
ID Q5LX88 SILPO PRELIMINARY; PRT; 419 AA.
AC Q5LX88;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Molybdopterin biosynthesis protein MoeA.
GN Name:moeA; OrderedLocusNames:SPO0310;
OS Silicibacter pomeroyi
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSB-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Kline R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA Rahe B., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment."
RL Nature 432:910-913(2004).
DR EMBL; CP000031; AAV93628.1; -; Genomic_DNA.
DR GO; GO:0006777; P:Mo-molybdopterin cofactor biosynthesis; IEA.
DR InterPro; IPR001453; MoCF_bios.
DR InterPro; IPR008284; MoCF_biosynth.
DR InterPro; IPR012086; MoeA.
DR InterPro; IPR005111; MoeA_C.
DR InterPro; IPR005110; MoeA_N.
DR Pfam; PF00994; MoCF_biosynth; 1.
DR Pfam; PF03454; MoeA_C; 1.
DR Pfam; PF03453; MoeA_N; 1.
DR PIRSF; PIRSF004870; MoeA; 1.
DR ProDom; PD002460; MoCF_biosynth; 1.
DR TIGRFAMs; TIGR00177; molyb_syn; 1.
DR PROSITE; PS01079; MoCF_BIOSYNTHESIS_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 419 AA; 43704 MW; BDE85DEB84CA0F4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 23
DB 51 LGGVLA 57
|||||

RESULT 595
Q4NJ12_9MICC
ID Q4NJ12_9MICC PRELIMINARY; PRT; 421 AA.
AC Q4NJ12;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Low temperature requirement A.
ORFNames=ArthDRAFT_3751;

```

```

OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-ORN);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHG01000002; EAL97884.1; -; Genomic_DNA.
SQ SEQUENCE 421 AA; 45516 MW; EB794B4FB51CLAID CRC64;

Query Match 5.9%; Score 7; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
DB 358 GVLAALA 364
|||||

RESULT 596
Q4HA72_9DEIO
ID Q4HA72_9DEIO PRELIMINARY; PRT; 421 AA.
AC Q4HA72;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Na+/H+ antiporter NhaA precursor.
GN ORFNames=DgeDRAFT_1086;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORN);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000004; EAL82984.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL 1 18 Potential.
SQ SEQUENCE 421 AA; 44042 MW; 1BB8993FA04E882B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 421;

```

Best Local Similarity 100.0%; Pred. No. 5.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 402 VLAALAA 408
|||||

RESULT 597
Q4UK37 RICFE PRELIMINARY; PRT; 421 AA.
ID Q4UK37
AC Q4UK37
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE AmpG protein.
GN Name=ampG3; OrderedLocusNames=RF_1247;
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia; spotted fever group.
OX NCBI_TaxID=42862;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=URRWXCal2;
RX PubMed=15984913; DOI=10.1371/journal.pbio.0030248;
RA Ogata H., Kenest P., Audic S., Robert C., Blanc G., Fournier P.-E.,
RA Parinello H., Claverie J.-M., Raoult D.;
RT "The genome sequence of Rickettsia felis identifies the first putative
RL conjugative plasmid in an obligate intracellular parasite.";
RL PLoS Biol. 3:E248-E248(2005).
DR EMBL; CP000053; AAY62098.1; -, Genomic_DNA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS0850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 421 AA; 47100 MW; 8761E027A61A1E5F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
Db 283 LGGVLA 289
|||||

RESULT 598
Q7MAC1 WOLSU PRELIMINARY; PRT; 421 AA.
AC Q7MAC1;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE UDP-N-ACETYLGLUCOSAMINE 1-CARBOXYVINYLTRANSFERASE ;MURA
DE (EC 2.5.1.7).
GN Name=MURA; OrderedLocusNames=WS0342;
OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteriaceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSMZ 1740;
RX MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.1932838100;
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
DR EMBL; BX571657; CAE09492.1; -, Genomic_DNA.
DR HSSP; P33038; 1DIG.
GO; GO:0016740; P:transferase activity; IEA.

DR GO; GO:0008760; F:UDP-N-acetylglucosamine 1-carboxyvinyltrans. .; IEA.
DR GO; GO:0019277; P:UDP-N-acetylglucosamine biosynthesis; IEA.
DR InterPro; IPR005750; AGlu_Tran_MurA.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR TIGRFAMs; TIGR01072; mura; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 421 AA; 45097 MW; 9C8D8FECE98B386B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 380 VLAALAA 386
|||||

RESULT 599
Q63MQ6 BURPS PRELIMINARY; PRT; 423 AA.
ID Q63MQ6
AC Q63MQ6
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=BPS0594;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K95243;
RX PubMed=15377794; DOI=10.1073/pnas.04033020101;
RA Holden M.T.G., Tittball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., DeShazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songsivilai S., Stevens K., Tumapa S., Vesaratchaveit M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571966; CAH38051.1; -, Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002453; Beta_cubulin.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS00228; TUBULIN_B_AUTOREG; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 423 AA; 43973 MW; 87C6220E032038DF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 386 VLAALAA 392
|||||

RESULT 600
Q4NFE9 9M1CC PRELIMINARY; PRT; 424 AA.
ID Q4NFE9 9M1CC PRELIMINARY;
AC Q4NFE9;
DT 13-SEP-2005 (TRENBLrel. 31, Created)

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Copper resistance protein CopC precursor.
 GN ORFNames=ArthDRAPT_1980;
 OS Arthrobacter sp. FB24.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcineae; Micrococcaceae; Arthrobacter.
 OX NCBI_TaxID=290399;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FB24;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israli S., Pitluck S., Richardson P., Richardson P.,
 RA "Sequencing of the draft genome assembly of Arthrobacter sp. FB24."
 RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=FB24;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24."
 RT Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AARG01000007; EAL95846.1; -; Genomic_DNA.
 KW Signal.
 FT SIGNAL
 FT SIGNAL
 SQ SEQUENCE 424 AA; 42524 MW; 1DCEA43032D60D8 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 424;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLAALAA 26
 DB 40 VLAALAA 46
 |||||
 RESULT 601
 Q9RLJ0_PSEAB
 ID Q9RLJ0_PSEAB PRELIMINARY; PRT; 426 AA.
 AC Q9RLJ0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE Putative signal transducer protein.
 DE CzcC protein.
 GN Name=czcC;
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CMG103;
 RK MEDLINE=20035746; PubMed=10570969; DOI=10.1016/S0378-1119(99)00349-2;
 RA Hassan M., van der Lelie D., Springael D., Romling N., Ahmed N.,
 RA "Identification of a gene cluster, czc, involved in cadmium and zinc
 RT resistance in Pseudomonas aeruginosa."
 RL Gene 238:417-425(1999).
 DR EMBL; Y14018; CAB56469.1; -; Genomic_DNA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003423; OEP.
 DR Pfam; PF02321; OEP; 2.
 DR PFam; PF02321; OEP; 2.
 SQ SEQUENCE 426 AA; 46292 MW; 6344F0E27FEEF7A CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 426;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ELGGKPA 47
 DB 104 ELGGKPA 110
 |||||
 RESULT 602
 P95363_NEIGO
 ID P95363_NEIGO PRELIMINARY; PRT; 427 AA.
 AC P95363;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Glutamine synthetase (glnA).
 OS Neisseria gonorrhoeae.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=485;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MS11;
 RK MEDLINE=97206157; PubMed=9157250;
 RA Zhou J., Bowler L.D., Spratt B.G.;
 RT "Interspecies recombination, and phylogenetic distortions, within the
 RT glutamine synthetase and shikimate dehydrogenase genes of Neisseria
 RT meningitidis and commensal Neisseria species."
 RL Mol. Microbiol. 23:799-812(1997).
 DR EMBL; U82701; AAC44903.1; -; Genomic_DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR004752; AmpG_permease.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR TIGRFAMs; TIGR00901; 2A0125; 1.
 SQ SEQUENCE 427 AA; 46349 MW; ACC491366B0A5324 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 427;
 Best Local Similarity 100.0%; Pred. No. 5.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAA 27
 DB 95 LAALAA 101
 |||||
 RESULT 603
 Q5F6G0_NEIG1
 ID Q5F6G0_NEIG1 PRELIMINARY; PRT; 427 AA.
 AC Q5F6G0;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Putative signal transducer protein.
 GN OrderedLocusNames=NGO1599;
 OS Neisseria gonorrhoeae (strain ATCC 700825 / FA 1090).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=242231;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RA Lewis L.A., Gillaspay A.F., McLaughlin R.E., Gipson M., Ducey T.F.,
 RA Ombey T., Hartman K., Nydick C., Carson M.B., Vaughn J., Thomson C.,
 RA Song L., Lin S., Yuan X., Najjar F., Zhan H., Ren Q., Zhu H., Qi S.,
 RA Kenton S.M., Lai H., White J.D., Clifton S., Roe B.A., Dyer D.W.;
 RT "The complete genome sequence of Neisseria gonorrhoeae."
 RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE004969; AAW90227.1; -; Genomic_DNA.
 DR InterPro; IPR004752; AmpG_permease.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR TIGRFAMs; TIGR00901; 2A0125; 1.
 KW Complete proteome.
 SQ SEQUENCE 427 AA; 46381 MW; 6113AC1E24E14632 CRC64;

```

Query Match          5.9%; Score 7; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 27
    |||||
DB 95 LAALAA 101

RESULT 604
ID Q9JSU7 NEIMA PRELIMINARY; PRT; 427 AA.
AC Q9JSU7_
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative integral membrane signal transducer protein.
GN OrderedLocusNames=NWA2127;
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Z2491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919; DOI=10.1038/35006655;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.M.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K.L., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis Z2491.";
RL Nature 404:502-506 (2000).
DR EMBL; AL162758; CAB85339.1; -; Genomic_DNA.
DR PIR; D81784; D81784.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004752; AMPG_Permease.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1.1.
DR TIGRFAMs; TIGR00901; 2A0125; 1.
KW Complete proteome.
SQ SEQUENCE 427 AA; 46310 MW; C4E79BAC56C5F2C3 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 427;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 27
    |||||
DB 95 LAALAA 101

RESULT 605
ID DAMX ECO57 STANDARD; PRT; 428 AA.
AC Q8X826;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DamX protein.
GN Name=damX; OrderedLocusNames=z4741, EC84230;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

Query Match          5.9%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QQAVIEP 102
    |||||
DB 259 QQAVIEP 265

RESULT 606
ID DAMX ECOLI STANDARD; PRT; 428 AA.
AC P11557; P76687;
DT 01-OCT-1989 (Rel. 12, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DamX protein.
GN Name=damX; OrderedLocusNames=b3388;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RX MEDLINE=89364696; PubMed=2549371; DOI=10.1007/BF00330946;
RA Janczyk P., Hines R., Smith D.W.;
RT "The Escherichia coli dam gene is expressed as a distal gene of a new
RT operon.";
RL Mol. Gen. Genet. 217:85-96 (1989).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=95327050; PubMed=7603433; DOI=10.1007/BF00290345;
RA Lyngstadaas A., Lobner-Olesen A., Boye E.;
RT "Characterization of three genes in the dam-containing operon of
RT Escherichia coli.";
RL Mol. Gen. Genet. 247:546-554 (1995).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

```

```

RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grofbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001).
CC -!- FUNCTION: Directly or indirectly interferes with cell division.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE005174; AAG58488.1; -; Genomic DNA.
CC EMBL; BA000007; BAB37653.1; -; Genomic_DNA.
CC PIR; D86003; D86003.
CC PIR; F91157; F91157.
CC Cell cycle; Cell division; Complete proteome; Transmembrane.
CC TRANSMEM 104 124 Potential.
CC SEQUENCE 428 AA; 46104 MW; 50D4A9527951058E CRC64;

Query Match          5.9%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QQAVIEP 102
    |||||
DB 259 QQAVIEP 265

RESULT 606
ID DAMX ECOLI STANDARD; PRT; 428 AA.
AC P11557; P76687;
DT 01-OCT-1989 (Rel. 12, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE DamX protein.
GN Name=damX; OrderedLocusNames=b3388;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12;
RX MEDLINE=89364696; PubMed=2549371; DOI=10.1007/BF00330946;
RA Janczyk P., Hines R., Smith D.W.;
RT "The Escherichia coli dam gene is expressed as a distal gene of a new
RT operon.";
RL Mol. Gen. Genet. 217:85-96 (1989).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=95327050; PubMed=7603433; DOI=10.1007/BF00290345;
RA Lyngstadaas A., Lobner-Olesen A., Boye E.;
RT "Characterization of three genes in the dam-containing operon of
RT Escherichia coli.";
RL Mol. Gen. Genet. 247:546-554 (1995).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

```

```
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R.; Plunkett G. III; Bloch C.A.; Perna N.T.; Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
CC -!- FUNCTION: Directly or indirectly interferes with cell division.
CC -!- INTERACTION:
CC P77310:nifu; NBEsp1; IntAct=EBI-561829, EBI-561646;
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; X15162; CAA33253.1; -; Genomic DNA.
CC EMBL; Z19601; CAA79667.1; -; Genomic DNA.
CC EMBL; U18997; AAA58185.1; -; Genomic DNA.
CC EMBL; U00096; AAC76413.1; -; Genomic DNA.
CC PIR; G65133; Q4ECAD.
CC IntAct; P11557; -.
CC EcoBASE; EBI1170; -.
CC EcoGene; EBI1183; damX
CC Cell cycle; Cell division; Complete proteome; Transmembrane.
CC TRANSMEM 104 124 Potential.
CC FT CONFLICT 6 6 P -> T (in Ref. 1 and 2).
CC FT CONFLICT 31 31 R -> P (in Ref. 1 and 2).
CC FT CONFLICT 86 91 RRPKR -> LRVYS (in Ref. 1 and 2).
CC FT CONFLICT 233 233 V -> L (in Ref. 2).
CC FT CONFLICT 310 310 A -> V (in Ref. 1 and 2).
CC FT CONFLICT 366 366 K -> N (in Ref. 2).
CC FT CONFLICT 382 382 G -> V (in Ref. 1 and 2).
CC SEQUENCE 428 AA; 46162 MW; E3376F88383BE94A CRC64;

Query Match 5.9%; Score 7; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 QQAVIEP 102
Db 259 QQAVIEP 265

RESULT 607
ID Q8PCV8_ECOL6 PRELIMINARY; PRT; 428 AA.
AC Q8PCV8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2004 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DamX protein.
GN Name=damX; OrderedLocNames=c4158;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O6.H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.25229799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Rosch P.,
RA Raiko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.P., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RL EMBL; A016768; AAN82596.1; -; Genomic DNA.
KW Complete proteome.
SQ SEQUENCE 428 AA; 46031 MW; 101EBD38CE57A902 CRC64;
```

```
Query Match 5.9%; Score 7; DB 2; Length 428;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 QQAVIEP 102
Db 259 QQAVIEP 265

RESULT 608
ID Q4NGI8_9MICC PRELIMINARY; PRT; 429 AA.
AC Q4NGI8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Major facilitator superfamily.
GN ORFNames=ArthDRAFT_1762;
OS Arthrobacter sp. FB24.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococciaceae; Micrococaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Terani S., Pitluck S., Richardson P.,
RT "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAH01000006; EAL96130.1; -; Genomic DNA.
SQ SEQUENCE 429 AA; 44368 MW; D3A676B114DFA445 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 351 GVLAALA 357

RESULT 609
ID Q7VVD4_BORPE PRELIMINARY; PRT; 429 AA.
AC Q7VVD4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative membrane transport protein.
GN OrderedLocNames=BP2742;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tohama 1 / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
```

RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640419; CAE43017.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; F:transporter activity; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
DR Complete proteome.
KW SEQUENCE 429 AA; 45627 MW; 9B18137E1B41F4F4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 164 VLAALAA 170
|||||

RESULT 610
Q7WKS4 BORBR PRELIMINARY; PRT; 429 AA.
AC Q7WKS4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=BB2030;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
EX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640443; CAE32527.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; F:transporter activity; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
DR Complete proteome.
KW SEQUENCE 429 AA; 45657 MW; 93574C5F46FED5AA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 164 VLAALAA 170
|||||

RESULT 610
Q7WKS4 BORBR PRELIMINARY; PRT; 429 AA.
AC Q7WKS4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=BB2030;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
EX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640443; CAE32527.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; F:transporter activity; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS50850; MFS; 1.
DR Complete proteome.
KW SEQUENCE 429 AA; 45657 MW; 93574C5F46FED5AA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 164 VLAALAA 170
|||||

RESULT 611
Q4XNNO_9DELT PRELIMINARY; PRT; 430 AA.
AC Q4XNNO;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AdhDRAFT_3291;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Iarani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000005; EAL80351.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 430 AA; 43851 MW; C1E273EBA90ECA36 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
Db 286 LLGGVLA 292
|||||

RESULT 612
Q83JA3_SHIFL PRELIMINARY; PRT; 430 AA.
AC Q83JA3; Q7EYS5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative membrane protein.
GN Name=damX; OrderedLocusNames=S4356, SF3406;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

Query Match 5.9%; Score 7; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 164 VLAALAA 170
|||||

RESULT 611
Q4XNNO_9DELT PRELIMINARY; PRT; 430 AA.
AC Q4XNNO;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AdhDRAFT_3291;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Iarani S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHD01000005; EAL80351.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 430 AA; 43851 MW; C1E273EBA90ECA36 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
Db 286 LLGGVLA 292
|||||

RESULT 612
Q83JA3_SHIFL PRELIMINARY; PRT; 430 AA.
AC Q83JA3; Q7EYS5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative membrane protein.
GN Name=damX; OrderedLocusNames=S4356, SF3406;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

DR TIGR; NMB0854; --; 1.
 DR HAMAP; MP_00127; --; 1.
 DR InterPro; IPR004154; anticodon_bd.
 DR InterPro; IPR004516; Hias.
 DR InterPro; IPR002314; tRNA-synt 2b.
 DR InterPro; IPR006195; tRNA ligase II.
 DR Pfam; PF03129; HGTP anticodon; 1.
 DR Pfam; PF00587; tRNA-synt 2b; 1.
 DR PIRSF; PIRSF001549; His-tRNA synth; 1.
 DR TIGRFAMs; TIGR00442; hias; 1.
 DR PROSITE; PS0862; AA tRNA LIGASE II; 1.
 KW Aminoacyl-tRNA synthetase; ATP-binding; Complete proteome; Ligase;
 KW Nucleotide-binding; Protein biosynthesis.
 SQ SEQUENCE 431 AA; 48465 MW; 63602B583A27DF12 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ELGGKPA 47
 |||||
 Db 302 ELGGKPA 308

RESULT 615
 ID Q5F9G8_NEIG1 PRELIMINARY; PRT; 431 AA.
 AC Q5F9G8;
 DT 10-MAY-2005 (TRENBLrel. 30, Created)
 DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
 DE Putative histidyl-tRNA synthetase (EC 6.1.1.21).
 GN OrderedLocusNames=NGO0426;
 OS Neisseria gonorrhoeae (strain ATCC 700825 / FA 1090).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 CC Neisseriaceae; Neisseria.
 OX NCBI_TaxID=242231;
 [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RA Lewis L.A., Gillaspay A.F., McLaughlin R.E., Gipson M., Ducey T.F.,
 RA Ombey T., Hartman K., Nydick C., Carson M.B., Vaughn J., Thomson C.,
 RA Song L., Lin S., Yuan X., Najjar F., Zhan M., Ren Q., Zhu H., Qi S.,
 RA Kenton S.M., Lai H., White J.D., Clifton S., Roe B.A., Dyer D.W.;
 RT "The complete genome sequence of Neisseria gonorrhoeae."
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 CC SUBMITTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
 diphosphate + L-histidyl-tRNA(His).
 CC - SUBUNIT: Homodimer (by similarity).
 CC - SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC - SIMILARITY: Belongs to the class-II aminoacyl-tRNA synthetase family.
 DR EMBL; AE004969; AAW89169.1; -; Genomic_DNA.
 DR GO; GO:0005737; C:cytoplasm; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004821; F:histidine-tRNA ligase activity; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0006427; P:histidyl-tRNA aminoacylation; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR004154; anticodon_bd.
 DR InterPro; IPR004516; Hias.
 DR InterPro; IPR002314; tRNA-synt 2b.
 DR InterPro; IPR006195; tRNA ligase II.
 DR Pfam; PF03129; HGTP anticodon; 1.
 DR Pfam; PF00587; tRNA-synt 2b; 1.
 DR PIRSF; PIRSF001549; His-tRNA synth; 1.
 DR TIGRFAMs; TIGR00442; hias; 1.
 DR PROSITE; PS0862; AA tRNA LIGASE II; 1.
 KW ATP-binding; Aminoacyl-tRNA synthetase; Complete proteome; Ligase;
 KW Nucleotide-binding; Protein biosynthesis.
 SQ SEQUENCE 431 AA; 48342 MW; 97DE6F818D1B259 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 ELGGKPA 47
 |||||
 Db 302 ELGGKPA 308

RESULT 616
 ID Q7XTK3_ORYSA PRELIMINARY; PRT; 435 AA.
 AC Q7XTK3;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DE OSJNBA0020P07.1 protein.
 GN Name=OSJNBA0020P07.1;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 EX PubMed=12447439; DOI=10.1038/nature01183;
 RA Feng Q., Zhang Y., Hao P., Wang S., Fu G., Huang Y., Li Y., Zhu J.,
 RA Liu Y., Hu X., Jia P., Zhang Y., Zhao Q., Ying K., Yu S., Tang Y.,
 RA Weng Q., Zhang L., Lu Y., Mu J., Lu Y., Zhang L.S., Yu S., Fan D.,
 RA Liu X., Lu T., Li C., Wu Y., Sun T., Lei H., Li T., Hu H., Guan J.,
 RA Wu M., Zhang R., Zhou B., Chen Z., Chen L., Jin Z., Wang R., Yin H.,
 RA Cai Z., Ren S., Lv G., Gu W., Zhu G., Tu Y., Jia J., Zhang Y.,
 RA Chen J., Kang H., Chen X., Shao C., Sun Y., Hu Q., Zhang X., Zhang W.,
 RA Wang L., Ding C., Sheng H., Gu J., Chen S., Ni L., Zhu F., Chen W.,
 RA Lan L., Lai Y., Cheng Z., Gu M., Jiang J., Li J., Hong G., Xue Y.,
 RA Han B.;
 RT "Sequence and analysis of rice chromosome 4."
 RL Nature 420:316-320(2002).
 DR EMBL; AL606450; CAE01284.1; -; Genomic_DNA.
 DR Gramene; Q7XTK3; -;
 DR GO; GO:0004040; F:amidase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000120; Amidase.
 DR PANTHER; PTHR11895; Amidase; 1.
 DR Pfam; PF01425; Amidase; 1.
 DR PROSITE; PS00571; AMIDASES; 1.
 SQ SEQUENCE 435 AA; 46179 MW; 68962A7745ED2F45 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 435;
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 |||||
 Db 69 VLAALAA 75

RESULT 617
 ID Q7JUN2_MYCPA PRELIMINARY; PRT; 435 AA.
 AC Q7JUN2;
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DE OSJNBA0020P07.1 protein.
 GN Name=aroA; OrderedLocusNames=MAP3334;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 OX NCBI_TaxID=1770;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=k10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;

Query Match 5.9%; Score 7; DB 2; Length 445;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
 DB 173 LLGGVLA 179

RESULT 621

Q5H4B3 XANOR
 ID Q5H4B3_XANOR PRELIMINARY; PRT; 449 AA.
 AC Q5H4B3;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DE H+ translocating pyrophosphate synthase.
 GN OrderedLocusNames=X000954;
 OS Xanthomonas oryzae (pv. oryzae).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 CC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=64187;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=KACC10331 / KX085;
 RX PubMed=15673718; DOI=10.1093/nar/gki206;
 RA Lee B.-M., Park Y.-J., Park D.-S., Kang H.-W., Kim J.-G., Song E.-S.,
 RA Park I.-C., Yoon U.-H., Hahn J.-H., Koo B.-S., Lee G.-B., Kim H.,
 RA Park H.-S., Yoon K.-O., Kim J.-H., Jung C.-H., Koh N.-H., Seo J.-S.,
 RA Go S.-J.;
 RT "The genome sequence of Xanthomonas oryzae pathovar oryzae KACC10331,
 RT the bacterial blight pathogen of rice."
 RL Nucleic Acids Res. 33:577-586(2005).
 DR EMBL; AB013598; AAW74208.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0009678; F:hydrogen-translocating pyrophosphatase acti. .; IEA.
 DR GO; GO:0004427; F:inorganic diphosphatase activity; IEA.
 DR GO; GO:0015992; F:proton transport; IEA.
 DR InterPro; IPR004131; H.PPase.
 DR Pfam; PF03030; H.PPase; 1.
 KW Complete proteome.
 SQ SEQUENCE 449 AA; 46230 MW; DC40117B8D8F705E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 449;
 Best Local Similarity 100.0%; Pred. No. 5.4e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLALA 25
 DB 73 GVLALA 79

RESULT 622

AROA_MYCBO
 ID AROA_MYCBO STANDARD; PRT; 450 AA.
 AC Q7TWY4;
 DT 10-MAY-2005 (Rel. 47, Created)
 DT 10-MAY-2005 (Rel. 47, Last sequence update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 GN Name=aroA; OrderedLocusNames=MB3256;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 CC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eiglmeyer K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 CC -|- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -|- PATHWAY: Metabolic intermediate biosynthesis; chorismate
 CC biosynthesis; chorismate from D-erythrose 4-phosphate and PEP;
 CC step 6.
 CC -|- PATHWAY: Context: Aromatic amino acids biosynthesis.
 CC -|- SUBUNIT: Monomer (By similarity).
 CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -|- SIMILARITY: Belongs to the EPSP synthase family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

CC EMBL; BX248345; CAD95348.1; -; Genomic_DNA.

DR HSSP; P07638; IG6S.

DR HAMAP; MF_00210; -; 1.

DR InterPro; IPR006264; AroA_Ctransf.

DR InterPro; IPR001986; EPSP_synth.

DR Pfam; PF00275; EPSP_synthase; 1.

DR ProDom; PD001867; EPSP_synth; 1.

DR TIGRFAMs; TIGR01356; aroA; 1.

DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.

DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.

KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;

SQ Complete proteome; Transferase.

Query Match 5.9%; Score 7; DB 1; Length 450;

Best Local Similarity 100.0%; Pred. No. 5.5e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26

DB 31 VLAALAA 37

RESULT 623

AROA_MYCTU
 ID AROA_MYCTU STANDARD; PRT; 450 AA.
 AC P22487;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
 GN Name=aroA; OrderedLocusNames=Rv3227, Mt3324; ORFNames=WTc20B11.02;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 CC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 CC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37Rv;
 RX MEDLINE=91072223; PubMed=2123856;
 RA Garbe T., Jones C., Charles I.G., Dougan G., Young D.;
 RT "Cloning and characterization of the aroA gene from Mycobacterium
 RL J. Bacteriol. 172:6774-6782(1990).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III,

Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornaby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
 Nature 393:537-544 (1998).
 [3]
 NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=CDC 1551 / Oehkosh;
 RC MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J.D., DeBoy R.T., Dodson R.J., Gwin M.L., Haft D.H., Hickey E.K., Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M., Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
 "Whole-genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";
 J. Bacteriol. 184:5479-5490 (2002).
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate = phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
 CC -1- PATHWAY: Metabolic intermediate biosynthesis; chorismate biosynthesis; chorismate from D-erythrose 4-phosphate and PEP; step 6.
 CC -1- PATHWAY: Context: Aromatic amino acids biosynthesis.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: Belongs to the EPSP synthase family.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

 DR EMBL; X52269; CAA36510.1; -; Genomic DNA.
 DR EMBL; M62708; AAA25356.1; -; Genomic DNA.
 DR EMBL; BX842582; CAB08328.1; -; Genomic DNA.
 DR EMBL; AE000516; AAK47667.1; -; Genomic DNA.
 DR PIR; E70590; E70590.
 DR HSP; P07638; 1G68.
 DR TIGR; MT3324; -.
 DR TubercuList; RV3227; -.
 DR HAMAP; MF_00210; -; 1.
 DR InterPro; IPR006264; AroA Ctransf.
 DR InterPro; IPR001986; EPSP_synth.
 DR Pfam; PF00275; EPSP_synthase; 1.
 DR ProDom; PD001867; EPSP_synth; 1.
 DR TIGRFAMs; TIGR01356; aroA; 1.
 DR PROSITE; PS00104; EPSP SYNTHASE 1; 1.
 DR PROSITE; PS00885; EPSP SYNTHASE 2; 1.
 KW Amino-acid biosynthesis; Aromatic amino acid biosynthesis;
 KW Complete proteome; Transferase.
 SQ SEQUENCE 450 AA; 46426 MW; 27BB86F9412A07D5 CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 450;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 20 VLAALAA 26
 Db 31 VLAALAA 37
 RESULT 624
 Q603J2_METCA
 ID Q603J2_METCA PRELIMINARY; PRT; 450 AA.
 AC Q603J2;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DR 4Fe-4S binding domain protein.
 GN OrderedLocuNames=MCA2813;
 OS Methylococcus capsulatus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
 OC Methylococcaceae; Methylococcus.
 OX NCBI_TaxID=414;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Bath / NCIMB 11132;
 RX PubMed=15383840; DOI=10.1371/journal.pbio.0020303;
 RA Ward N.L., Larsen O., Sakwa J., Bruseeth L., Khouri H.M., Durkin A.S., Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E., Methe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E., Ravel J., Tettelin H., Ren Q., Read T.D., DeBoy R.T., Seshadri R., Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J., Grindhaug S.H., Holt I.E., Eidhammer I., Jonassen I., Vanaken S., Uterback T.R., Feldblyum T.V., Fraser C.M., Lillehaug J.R., Eisen J.A.;
 "Genomic insights into methanotrophy: the complete genome sequence of Methylococcus capsulatus (Bath).";
 PLoS Biol. 2:1616-1628 (2004).
 DR EMBL; AE017282; AAU91052.1; -; Genomic DNA.
 DR TIGR; MCA2813; -.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0005506; F:iron ion binding; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001450; 4Fe4S_Fe_S_bd.
 DR Pfam; PF00037; Fer4; 1.
 DR PRINTS; PR00353; 4FE4SPROXIN
 DR PROSITE; PS00198; 4FE4S_FEREDOXIN; 1.
 KW 4Fe-4S; Complete proteome; Electron transport; Iron; Iron-sulfur;
 KW Metal-binding; Transport.
 SQ SEQUENCE 450 AA; 50406 MW; 2829B5934D037E78 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 450;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 25 AAYCLSV 31
 Db 276 AAYCLSV 282
 RESULT 625
 Q5E7W9_VIBF1
 ID Q5E7W9_VIBF1 PRELIMINARY; PRT; 451 AA.
 AC Q5E7W9;
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, Last sequence update)
 DE Mg(2+) transporter MgtE.
 GN OrderedLocuNames=VF0382;
 OS Vibrio fischeri (strain ATCC 700601 / ES114).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=312309;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15703294; DOI=10.1073/pnas.0409900102;
 RA Ruby E.G., Urbanowski M., Campbell J., Dunn A., Faini M., Gunsalus R., Lostroh P., Lupp C., McCann J., Millikan D., Schaefer A., Stabb E., Stevens A., Visick K., Whistler C., Greenberg E.P.;
 "Complete genome sequence of Vibrio fischeri: a symbiotic bacterium with pathogenic congeners.";
 Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009 (2005).
 DR EMBL; CP000020; AAW84877.1; -; Genomic DNA.
 DR InterPro; IPR000644; CBS.
 DR InterPro; IPR006667; MgtE integratmembr.
 DR InterPro; IPR006668; MgtE intracel.
 DR InterPro; IPR006669; MgtE_transporter.

```
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF01769; MgtE; 1.
DR Pfam; PF03448; MgtE_N; 1.
DR TIGRFAMs; TIGR00400; mgtE; 1.
KW Complete proteome.
SQ SEQUENCE 451 AA; 48885 MW; 3C51AD23746892CA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 296 VLAALAA 302

RESULT 626
Q8DE98 VIBVU PRELIMINARY; PRT; 451 AA.
AC Q8DE98;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Mg/Co/Ni transporter MgtE.
GN OrderedLocusNames=V10698;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR GO; GO:0008324; P: cation transporter activity; IEA.
DR GO; GO:0006812; P: cation transport; IEA.
DR InterPro; IPR00644; CBS.
DR InterPro; IPR006667; MgtE_integrmembr.
DR InterPro; IPR006668; MgtE_intracl.
DR InterPro; IPR006669; MgtE_transporter.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF01769; MgtE; 1.
DR Pfam; PF03448; MgtE_N; 1.
DR SMART; SM00116; CBS; 1.
DR TIGRFAMs; TIGR00400; mgtE; 1.
KW Complete proteome.
SQ SEQUENCE 451 AA; 48798 MW; 58AFC8A3EF37BF0D CRC64;

Query Match 5.9%; Score 7; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 296 VLAALAA 302

RESULT 627
Q87LD7 VIBPA PRELIMINARY; PRT; 451 AA.
AC Q87LD7;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Magnesium transporter.
GN OrderedLocusNames=VP2675;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;

[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RA MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yaenunga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae."
RL Lancet 361:743-749(2003).
DR EMBL; BA000031; BAC60938.1; -; Genomic DNA.
DR GO; GO:0008324; P: cation transporter activity; IEA.
DR GO; GO:0006812; P: cation transport; IEA.
DR InterPro; IPR006644; CBS.
DR InterPro; IPR006667; MgtE_integrmembr.
DR InterPro; IPR006668; MgtE_intracl.
DR InterPro; IPR006669; MgtE_transporter.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF01769; MgtE; 1.
DR Pfam; PF03448; MgtE_N; 1.
DR SMART; SM00116; CBS; 1.
DR TIGRFAMs; TIGR00400; mgtE; 1.
KW Complete proteome.
SQ SEQUENCE 451 AA; 48914 MW; D7A25878F475C3B2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 296 VLAALAA 302

RESULT 628
Q7MPC2 VIBVY PRELIMINARY; PRT; 451 AA.
AC Q7MPC2;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Magnesium transporter.
GN OrderedLocusNames=VV0442;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen."
RL Genome Res. 13:2577-2587(2003).
DR EMBL; BA000037; BAC93206.1; -; Genomic DNA.
DR GO; GO:0008324; P: cation transporter activity; IEA.
DR GO; GO:0006812; P: cation transport; IEA.
DR InterPro; IPR006644; CBS.
DR InterPro; IPR006667; MgtE_integrmembr.
DR InterPro; IPR006668; MgtE_intracl.
DR InterPro; IPR006669; MgtE_transporter.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF01769; MgtE; 1.
DR Pfam; PF03448; MgtE_N; 1.
DR TIGRFAMs; TIGR00400; mgtE; 1.
KW Complete proteome.
SQ SEQUENCE 451 AA; 48798 MW; 58AFC8A3EF37BF0D CRC64;

Query Match 5.9%; Score 7; DB 2; Length 451;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

Qy      20 VLAALAA 26
Db      296 VLAALAA 302
|||||

RESULT 629
Q6LMB1_PROPR PRELIMINARY;      PRT;      452 AA.
AC Q6LMB1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative magnesium transporter.
GN Name=NE1633; OrderedLocustNames=PPRA32260;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15746425; DOI=10.1126/science.1103341;
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N.,
RA Lauro F.M., Cestaro A., Malacrida G., Simonati B., Cannata N.,
RA Romualdi C., Bartlett D.H., Valle G.;
RT "Life at depth: Photobacterium profundum genome sequence and
RT expression analysis.";
RL Science 307:1459-1461(2005).
DR EMBL; CR378673; CAG21566.1; -; Genomic DNA.
DR GO; GO:0008324; P:cation transporter activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR006667; MgtE_integrmembr.
DR InterPro; IPR006668; MgtE_intracel.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF01769; MgtE; 1.
DR Pfam; PF03448; MgtE_N; 1.
DR TIGRFAMs; TIGR00400; mgtE; 1.
KW Complete proteome.
SQ SEQUENCE 452 AA; 49194 MW; 0B2A081AB925ACAA0 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      20 VLAALAA 26
Db      296 VLAALAA 302
|||||

RESULT 630
Q8RGRL_FUSNN PRELIMINARY;      PRT;      452 AA.
AC Q8RGRL;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Gluconate permease.
GN OrderedLocustNames=FN0225;
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RX DOI=10.1128/JB.184.7.2005-2018.2002;
RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Shattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,

```

```

RA Fonstein M., Kyripides N.C., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RL nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE009951; AAL94431.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0015128; F:gluconate transporter activity; IEA.
DR GO; GO:0015725; P:gluconate transport; IEA.
DR InterPro; IPR003474; Glcn transporter.
DR Pfam; PF02447; GntP_permease; 1.
DR TIGRFAMs; TIGR00791; gntP; 1.
KW Complete proteome.
SQ SEQUENCE 452 AA; 47782 MW; 3E94AA95076CFD59 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 452;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      20 VLAALAA 26
Db      389 VLAALAA 395
|||||

RESULT 631
Q9KP45_VIBCH PRELIMINARY;      PRT;      453 AA.
AC Q9KP45;
ID Q9KP45_VIBCH PRELIMINARY;      PRT;      453 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Magnesium transporter.
GN OrderedLocustNames=VC2534;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umavam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Niernan W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AR004322; AAF95675.1; -; Genomic_DNA.
DR PIR; C82066; C82066.
DR TIGR; VC2534; -.
DR GO; GO:0008324; P:cation transporter activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR InterPro; IPR000644; CBS.
DR InterPro; IPR006667; MgtE_integrmembr.
DR InterPro; IPR006668; MgtE_intracel.
DR InterPro; IPR006669; MgtE_transporter.
DR Pfam; PF00571; CBS; 1.
DR Pfam; PF01769; MgtE; 1.
DR Pfam; PF03448; MgtE_N; 1.
DR SMART; SM00116; CBS; 1.
DR TIGRFAMs; TIGR00400; mgtE; 1.
KW Complete proteome.
SQ SEQUENCE 453 AA; 49152 MW; 55353CA3CC590C3B CRC64;

Query Match      5.9%; Score 7; DB 2; Length 453;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      20 VLAALAA 26
Db      389 VLAALAA 395
|||||

```


RA Souza R.C., Steffens M.B.R., Steindel M., Teixeira S.R., Urmenyi T.,
 RA Vettore A., Wassem R., Zaha A., Simpson A.J.G.;
 RT "the complete genome sequence of *Chromobacterium violaceum* reveals
 RL remarkable and exploitable bacterial adaptability.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:11660-11665(2003).
 DR EMBL; AA059914.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008289; F:lipid binding; IEA.
 DR GO; GO:0005215; P:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003423; OSP.
 DR InterPro; IPR010131; RND_outer_NodT.
 DR Pfam; PF02321; OEP; 2.
 DR TIGRFAMs; TIGR01845; RND_outer_NodT; 1.
 KW Complete proteome; Lipoprotein.
 SQ SEQUENCE 459 AA; 48518 MW; 8662B74F63F78190 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 27
 Db 359 LAALAA 365
 |||||

RESULT 635
 Q7UAV7_SYNXPX
 ID Q7UAV7_SYNXPX PRELIMINARY; PRT; 459 AA.
 AC Q7UAV7;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative aldehyde dehydrogenase (EC 1.2.1.3).
 GN OrderedLocuNames=SYNW1956;
 OS *Synechococcus* sp. (strain WH8102).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=84588;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
 RA Palenik B., Brahamsha B., Laximer F.W., Land M.L., Hauser L.,
 RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarron J.,
 RA Paulsen I.T., Dufresne A., Partensky F., Webb E.A., Waterbury J.;
 RT "the genome of a motile marine *Synechococcus*.";
 RL Nature 424:1037-1042(2003).
 CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
 DR EMBL; BX569694; CAB08471.1; -; Genomic_DNA.
 DR HGSP; P11883; IAD3.
 DR GO; GO:0004029; P:aldehyde dehydrogenase (NAD) activity; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002086; Aldehyd_dehydrog.
 DR Pfam; PF00171; Aldehd; 1.
 DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
 KW Complete proteome; Oxidoreductase.
 SQ SEQUENCE 459 AA; 49782 MW; 9D5BF79E58CFD95 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 459;
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 Db 44 VLAALAA 50
 |||||

RESULT 636
 Q73Y31_MYCPA
 ID Q73Y31_MYCPA PRELIMINARY; PRT; 463 AA.
 AC Q73Y31;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocuNames=MAP2127;
 OS *Mycobacterium paratuberculosis*.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC *Mycobacterium avium* complex (MAC).
 OX NCBI_TaxID=1770;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=k10;
 RL Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB017234; AAS04444.1; -; Genomic_DNA.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR011251; Luciferase like.
 DR PRINTS; PR01415; ANKYRIN_BOX_N.
 DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 463 AA; 50551 MW; 3E1CC8466FD9E8DC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 463;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 Db 85 VLAALAA 91
 |||||

RESULT 637
 Q9PGM2_XYLFA
 ID Q9PGM2_XYLFA PRELIMINARY; PRT; 471 AA.
 AC Q9PGM2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE UDP-N-acetylmuramate-L-alanine ligase.
 GN OrderedLocuNames=Xf0276;
 OS *Xylella fastidiosa*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=9a5c;
 RX MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
 RA Colauto N.B., Colombo C., Costa P.F., Costa M.C.R., Costa-Neto C.M.,
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Pacinani A.P., Pereira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca B.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.B.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
 RA de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tuhako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,

RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";
 RL Nature 406:151-159(2000).
 DR EMBL; AE003881; AAF83089.1; -; Genomic DNA.
 DR PIR; C82825; C82825.
 DR HSSP; P45066; 1GOY.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0036874; F:ligase activity; IEA.
 DR GO; GO:0009058; F:biosynthesis; IEA.
 DR GO; GO:0007047; P:cell wall organization and biogenesis; IEA.
 DR InterPro; IPR002110; ANK.
 DR InterPro; IPR005757; Mpl.
 DR InterPro; IPR004101; Mur_ligase_C.
 DR InterPro; IPR000713; Mur_ligase_N.
 DR InterPro; IPR012237; UDP-NACM_Align.
 DR Pfam; PF01225; Mur_ligase; 1.
 DR Pfam; PF02875; Mur_ligase_C; 1.
 DR PIRSF; PIRSF01562; UDP-NACM_Align; 1.
 DR PRINTS; PR01415; ANKYRIN.
 DR TIGRFAMs; TIGR01081; mpl; 1.
 DR Complete proteome; Ligase.
 KW SEQUENCE 471 AA; 50775 MW; 587356E607975579 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 471;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 306 VLAALAA 312

RESULT 638
 Q7V8W3 PROMM PRELIMINARY; PRT; 471 AA.
 AC Q7V8W3_1
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Putative multidrug efflux transporter, MFS family.
 GN OrderedLocustNames=PWT0210;
 OS *Prochlorococcus marinus* (strain MIT 9313).
 OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
 OC *Prochlorococcus*.
 CX NCBI_TaxID=74547;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22825598; PubMed=12917642; DOI=10.1038/nature01947;
 RA Kocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
 RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
 RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;
 RT "Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic
 niche differentiation.";
 RL Nature 424:1042-1047(2003).
 DR EMBL; BX572095; CAE20385.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 KW Complete proteome.
 SQ SEQUENCE 471 AA; 50250 MW; A65D85A75D5753F3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 471;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
 |||||

Db 373 VLLGGVL 379
 RESULT 639
 Q8KH7 NOCAE PRELIMINARY; PRT; 473 AA.
 AC Q8KH7_1
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Putative antibiotic antiporter (Putative translocase protein)
 DE Putative integral membrane transporter.
 GN Name=rbmk; Synonyms=rebT;
 OS *Nocardia aerocolonigenes*.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Pseudonocardiaceae; Actinosynnemataceae; Lechevalieria.
 CX NCBI_TaxID=68170;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Hyun C.-G., Billign T., Liao J., Thorson J.S.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Nishizawa T., Aldrich C.C., Sherman D.H.;
 RT "Molecular Analysis of the Rebeccamycin L-Amino Acid Oxidase from
 Lechevalieria aerocolonigenes ATCC 39243.";
 RL J. Bacteriol. 187:2084-2092(2005).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=ATCC 39243;
 RA Sanchez C., Butovich I.A., Brana A.F., Rohr J., Mendez C., Salas J.A.;
 RT "The biosynthetic gene cluster for the antitumor rebeccamycin:
 characterization and generation of indolocarbazole derivatives.";
 RL Chem. Biol. 9:519-531(2002).
 DR EMBL; AF534707; AAN01217.1; -; Genomic DNA.
 DR EMBL; AB090952; BAC10683.1; -; Genomic DNA.
 DR EMBL; AJ414559; CAC93723.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0046677; P:response to antibiotic; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005829; Sug_transporter.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN 1.
 SQ SEQUENCE 473 AA; 47762 MW; C54E0955197B277E CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 473;
 Best Local Similarity 100.0%; Pred. No. 5.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
 |||||
 Db 153 VLLGGVL 159

RESULT 640
 Q5YV45 NOCAE PRELIMINARY; PRT; 479 AA.
 AC Q5YV45_1
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative transporter.
 GN OrderedLocustNames=nfa30990;
 OS *Nocardia farcinica*.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriaceae; Nocardiaceae; Nocardia.
 CX NCBI_TaxID=37329;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

```
RC STRAIN=IFM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Iehikawa J., Yanashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BA057946.1; -; Genomic_DNA.
GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR PROSITE; PS00850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 479 AA; 47941 MW; 1B756C254FC9DD50 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 154 VLLGGVL 160

RESULT 641
ID Q7U3Q4 SYNXP PRELIMINARY; PRT; 480 AA.
AC Q7U3Q4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=SYN2376;
OS Synchococcus sp. (strain WH8102).
OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.
OX NCBI_TaxID=8458;
RN [1]
RP MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;
RA Palenik B., Brahmeha B., Larimer F.W., Land M.L., Hauser L.,
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCaren J.,
RA Paulsen I.T., Dufrene A., Partensky F., Webb E.A., Waterbury J.;
RT "The genome of a motile marine Synchococcus.";
RL Nature 424:1037-1042(2003).
DR EMBL; BX569695; CA080891.1; -; Genomic_DNA.
DR InterPro; IPR009651; Alum.res.
DR Pfam; PF06838; Alum.res.; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 480 AA; 50684 MW; 578B147C47E41A4E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 94 VLAALAA 100

RESULT 642
ID Q62165 BURMA PRELIMINARY; PRT; 480 AA.
AC Q62165;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Drug resistance transporter, EmrB/QacA subfamily.
GN OrderedLocNames=BMA2023;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
```

```
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344; DOI=10.1073/pnas.0403306101;
RX PubMed=15377793; PubMed=15377794; DOI=10.1073/pnas.0403306101;
RA Nieman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.L., Roming C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.P.,
RA Madupu R., Mohamoud Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamblin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000010; AAU49554.1; -; Genomic_DNA.
DR TIGR; BMA2023; -;
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004638; Efflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR TIGRFAMS; TIGR00711; efflux_EmrB; 1.
DR PROSITE; PS00850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 480 AA; 48701 MW; 0A5P5E6B6ACDF1BE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 154 VLLGGVL 160

RESULT 643
ID Q63RF9 BURPS PRELIMINARY; PRT; 480 AA.
AC Q63RF9;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative membrane transport protein.
GN OrderedLocNames=BPSL2713;
OS Burkholderia pseudomallei (Pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales.
OX Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K36243; DOI=10.1073/pnas.0403302101;
RX PubMed=15377794; PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.P., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., Deshazer D.,
RA Feitwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchaveit M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis, Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH36721.1; -; Genomic_DNA.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
```

```

DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0046677; P:response to antibiotic; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004638; Eflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR011701; MFS_1.
DR Pfam; PF07690; MFS_1; 1.
DR TIGRFAMs; TIGR00711; eflux_EmrB; 1.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 480 AA; 48759 MW; 0BF91C0DF8385032 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 480;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 154 VLLGGVL 160

RESULT 644
Q75HF3_ORYSA PRELIMINARY; PRT; 483 AA.
AC Q75HF3;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNB0024F18.8.
GN Name=OSJNB0024F18.8;
OS Oryza sativa (japonica cultivar group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feidblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Buell R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135594; AAR89844.1; -; Genomic_DNA.
DR Gramene; Q75HF3; -.
DR Hypothetical protein.
SQ SEQUENCE 483 AA; 55035 MW; B496F5495AA42EA9 CRC64;

Query Match          5.9%; Score 7; DB 2; Length 483;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 228 GVLAALA 234

RESULT 645
MURE_RHILO STANDARD; PRT; 484 AA.
AC Q98KA8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE UDP-N-acetylmuramoyl-L-alanyl-D-glutamate--2,6-diaminopimelate ligase
DE (EC 6.3.2.13) (UDP-N-acetylmuramyl-tripeptide synthetase) (Meso-
DE diaminopimelate-adding enzyme) (UDP-MurNAC-tripeptide synthetase).
GN Name=murE; OrderedLocusNames=ml11560;

Rhizobium loti (Mesorhizobium loti).
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Phyllobacteriaceae; Mesorhizobium.
NCBI_TaxID=381;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=MAFF303099; PubMed=11214968;
RX MEDLINE=21082930; Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Kaneko T., Nakamura Y., Itoh K., Kikawa K., Kimura T.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Matsumoto M.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
CC -!- FUNCTION: Cell wall formation. Diaminopimelic acid adding enzyme
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
CC glutamate + meso-2,6-diaminoheptanedioate = ADP + phosphate + UDP-
CC N-acetylmuramoyl-L-alanyl-D-gamma-glutamyl-meso-2,6-diamino-
CC heptanedioate.
CC -!- PATHWAY: Peptidoglycan biosynthesis.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the murCDEF family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BA000012; BAB48906.1; -; Genomic_DNA.
DR HSSP; P22188; 1E8C.
DR HAMAP; MF_00208; 1.
DR InterPro; IPR004101; Mur_ligase_C.
DR InterPro; IPR000713; Mur_ligase_N.
DR InterPro; IPR005761; MurE.
DR InterPro; IPR012237; UDP-NacMg_Align.
DR Pfam; PF01225; Mur_ligase_1.
DR Pfam; PF02875; Mur_ligase_C; 1.
DR PIRSF; PIRSF001562; UDP-NacMg_Align; 1.
DR TIGRFAMs; TIGR01085; murE; 1.
KW ATP-binding; Cell cycle; Cell division; Cell shape; Cell wall;
KW Complete proteome; Ligase; Nucleotide-binding;
KW Peptidoglycan synthesis.
FT NP_BIND 108 114 ATP (Potential).
SQ SEQUENCE 484 AA; 50751 MW; 1F9652A301F13E54 CRC64;

Query Match          5.9%; Score 7; DB 1; Length 484;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 SADLEVT 10
Db 17 SADLEVT 23

RESULT 646
Q4NTY0_9DELT PRELIMINARY; PRT; 486 AA.
AC Q4NTY0;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Phospholipase D/Transphosphatidylase.
GN ORFNames=AdelDRAFT_2105;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophacteriaceae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGP);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hammon N., Ierani S., Pittluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL; AAHD01000016; EAL78990.1; -; Genomic DNA.
DR EMBL; AAHD01000016; 52141 MW; 093DEC0745C37936 CRC64;
SQ SEQUENCE 486 AA; 52141 MW; 093DEC0745C37936 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 162 VLAALAA 168

RESULT 647
Q7YVN3_9TRYP PRELIMINARY; PRT; 488 AA.
AC Q7YVN3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Expression site associated protein 5, putative (Expression site-
DE associated gene (ESAG) protein, putative).
GN ORFNames=Tb927.2.1920;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
[1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22789168; PubMed=12907728; DOI=10.1093/nar/gkg673;
RA El-Sayed N.M.A., Ghedin E., Song J., MacLeod A., Bringaud F.,
RA Larkin C., Wanless D., Peterson J., Hou L., Taylor S., Tweedie A.,
RA Biteau N., Khalak H.G., Lin X., Mason T., Hannick L., Caler E.,
RA Blandin G., Bartholomeu D., Simpson A.J., Kaul S., Zhao H., Pai G.,
RA Van Aken S., Utterback T., Haas B., Koo H.L., Umayam L., Suh B.,
RA Gerrard C., Leech V., Qi R., Zhou S., Schwartz D., Feldblyum T.,
RA Salzberg S., Tait A., Turner M.R., Ullu E., White O., Melville S.,
RA Adams M.D., Fraser C.M., Donelson J.E.;
RT "The sequence and analysis of Trypanosoma brucei chromosome II.";
RN Nucleic Acids Res. 31:4856-4863(2003).
[2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RG Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shallow J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RG El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;

RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK017168; AAQ15650.1; -; Genomic DNA.
DR EMBL; AC008031; AAQ19181.1; -; Genomic DNA.
SQ SEQUENCE 488 AA; 53133 MW; 0CAFA03F4C02CAB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DLEVTTS 12
| | | | |
Db 174 DLEVTTS 180

RESULT 648
Q6DHU0_BRARE PRELIMINARY; PRT; 488 AA.
AC Q6DHU0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Zgc:92064 protein.
GN ORFNames=zgc:92064;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D., Woxley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RA Director MGC Project;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; BC075877; AAH75877.1; -; mRNA.
DR ZFIN; ZDB-GENE-040718-74; zgc:92064.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase.
SQ SEQUENCE 488 AA; 54656 MW; D681B6495357E877 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 20 VLAALAA 26
Db 474 VLAALAA 480

RESULT 649
Q90ZZ8 BRARE PRELIMINARY; PRT; 488 AA.
AC Q90ZZ8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Aldehyde dehydrogenase (EC 1.2.1.3).
GN ORFNames=wu:fc06b11, zgc:92064;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RP NUCLEOTIDE SEQUENCE.
RA Pappa A., Tanguay R., Reimers M., Vasilou V.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the aldehyde dehydrogenase family.
DR EMBL; AF254954; AAK49120.1; -; mRNA.
DR HSSP; P11883; 1AD3.
DR Ensembl; ENSDARG0000032106; Danio rerio.
DR ZFIN; ZDB-GENE-040718-74; zgc:92064.
DR GO; GO:0004029; F:aldehyde dehydrogenase (NAD) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002086; Aldehyd_dehydrog.
DR Pfam; PF00171; Aldedh; 1.
DR PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; 1.
DR PROSITE; PS00687; ALDEHYDE_DEHYDR_GLU; 1.
KW Oxidoreductase.
SQ SEQUENCE 488 AA; 54674 MW; 04E9EB7B67195686 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 488;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 VLAALAA 26
Db 474 VLAALAA 480

RESULT 650
Q5B457 EMENI PRELIMINARY; PRT; 489 AA.
AC Q5B457;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AN4673.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=227321;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgaltier B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galleagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,

```

```

RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Maucelli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupbach R., Seanan S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Tamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RL "Genome Sequence of Aspergillus nidulans.";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC Preliminary data.
DR EMBL; AACD01000080; EAA60715.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 489 AA; 56193 MW; 6EECFDA53F60FAE2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 5.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 18 GGVLAAL 24
Db 207 GGVLAAL 213

RESULT 651
Q5LUT5_SILPO PRELIMINARY; PRT; 492 AA.
AC Q5LUT5;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Phosphate transporter family protein.
GN OrderedLocusNames=SP00967;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
[1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature03170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heldeberg J.F., Whitman W.B.,
RA Kiene R.P., Henriksen J.R., King G.M., Belas R., Fuqua C.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment.";
RL Nature 432:910-913(2004).
DR EMBL; CP000031; AAV94272.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005315; F:inorganic phosphate transporter activity; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR001204; Phos_transporter;
DR PANTHER; PTHR11101; Phos_transporter; 2.
DR Pfam; PF01384; PHO4; 1.
KW Complete proteome.
SQ SEQUENCE 492 AA; 51407 MW; 0FDEC92CA74425F4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 492;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 20 VLAALAA 26

```

|||||
45 VLAALAA 51

```

Db
RESULT 652
PDI2 LABEL STANDARD; PRT; 493 AA.
AC Q1770;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Protein disulfide-isomerase 2 precursor (EC 5.3.4.1) (PDI 1) (Prolyl
DE 4-hydroxylase beta subunit).
GN Name=pdi-2; ORFName=C07A12.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE
RX MEDLINE=96332462; PubMed=8760355;
RA Veljola J., Annunen P., Kolvinen P., Page A.P., Pihlajaniemi T.,
RA Kivirikko K.I.;
RT "Baculovirus expression of two protein disulfide isomerase isoforms
RT from Caenorhabditis elegans and characterization of prolyl 4-
RT hydroxylases containing one of these polypeptides as their beta
RT subunit.";
RL Blochem. J. 317:721-729(1996).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [3]
RP FUNCTION.
RX MEDLINE=20266296; PubMed=10805750;
RX DOI=10.1126/SCIENCE.111.4084-4093.2000;
RA Winter A.D., Page A.P.;
RT "Prolyl 4-hydroxylase is an essential procollagen-modifying enzyme
RT required for exoskeleton formation and the maintenance of body shape
RT in the nematode Caenorhabditis elegans.";
RL J. Biol. Chem. 277:29187-29196(2002).
RL Mol. Cell. Biol. 20:4084-4093(2000).
RN [4]
RP SUBUNIT.
RX MEDLINE=22151124; PubMed=12036960; DOI=10.1074/jbc.M203824200;
RA Myllyharju J., Kukkola L., Winter A.D., Page A.P.;
RT "The exoskeleton collagens in Caenorhabditis elegans are modified by
RT prolyl 4-hydroxylases with unique combinations of subunits.";
RL J. Biol. Chem. 277:29187-29196(2002).
CC -1- CATALYTIC ACTIVITY: Catalyzes the rearrangement of -S-S- bonds in
CC proteins.
CC -1- CATALYTIC ACTIVITY: Procollagen L-proline + 2-oxoglutarate + O(2)
CC -1- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
CC Exist either as a phy-1(2)/pdi-2(2) tetramer, a phy-2(2)/pdi-2(2)
CC tetramer or as a phy-1/phy-2/pdi-2(2) tetramer.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -1- SIMILARITY: Belongs to the protein disulfide isomerase family.
CC -1- SIMILARITY: Contains 2 thioredoxin domains.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U41542; AKG39152.1; -; Genomic_DNA.
CC PIR; T34092; T34092.
CC HSP; P07237; 1B7X.
CC Ensembl; C07A12.4; Caenorhabditis elegans.

```

```

DR WormBase; WBGene00003963; pdi-2.
DR WormPep; C07A12.4a; C03972.
DR InterPro; IPR005788; Disulphid_ism.
DR InterPro; IPR005792; Disulphide_ism.
DR InterPro; IPR006662; Thioered.
DR InterPro; IPR006663; Thioeredox dom2.
DR InterPro; IPR012336; Thioeredoxin-like.
DR InterPro; IPR012335; Thioeredoxin_fold.
DR Pfam; PF00085; Thioeredoxin; 3.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01130; ER_PDI_fam; 1.
DR TIGRFAMs; TIGR01126; pdi_dom; 2.
DR PROSITE; PS00194; THIOREDOXIN; 2.
KW Complete proteome; Endoplasmic reticulum; Isomerase;
KW Redox-active center; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 493 Protein disulfide-isomerase 2.
FT DOMAIN 24 127 Thioeredoxin 1.
FT DOMAIN 364 467 Thioeredoxin 2.
FT MOTIF 490 493 Prevents secretion from ER (Potential).
FT ACT_SITE 52 52 Nucleophile (By similarity).
FT ACT_SITE 55 55 Nucleophile (By similarity).
FT ACT_SITE 393 393 Nucleophile (By similarity).
FT ACT_SITE 396 396 Nucleophile (By similarity).
FT SITE 53 53 Contributes to redox potential value (By similarity).
FT SITE 54 54 Contributes to redox potential value (By similarity).
FT SITE 116 116 Lowers pKa of C-terminal Cys of first active site (By similarity).
FT SITE 394 394 Contributes to redox potential value (By similarity).
FT SITE 395 395 Contributes to redox potential value (By similarity).
FT SITE 456 456 Lowers pKa of C-terminal Cys of second active site (By similarity).
FT DISULFID 52 55 Redox-active (By similarity).
FT DISULFID 393 396 Redox-active (By similarity).
SQ SEQUENCE 493 AA; 55152 MW; 41BC0C1185BFB85E CRC64;

Query Match 5.9%; Score 7; DB 1; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 81 QPFGKVL 87
|||
Db 277 QPFGKVL 283

RESULT 653
Q4USF6_DICVI
ID Q4USF6_DICVI PRELIMINARY; PRT; 493 AA.
AC Q4USF6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
OS Dictyocaulus viviparus (Bovine lungworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Dictyocaulidae; Dictyocaulinae; Dictyocaulus.
OC NCBI_TaxID=29172;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Wolken S., von Samson-Himmelstjerna G., Schnieder T.;
RT "Characterization of putative protein disulfide isomerases in the
RT bovine lungworm Dictyocaulus viviparus.";
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RW EMBL; DQ011044; AY333776.1; -; mRNA.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 54790 MW; D94AAD1F2BAEBDC1 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;

```

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
Db 277 QPKGKVL 283

RESULT 654
Q618U9 CAEBR
ID Q618U9 CAEBR PRELIMINARY; PRT; 493 AA.
AC Q618U9;
DT 25-OCT-2004 (TREMREL. 28, Created)
DT 25-OCT-2004 (TREMREL. 28, Last sequence update)
DE Hypothetical protein CBG14484.
GN Name=CBG14484;
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermidae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; CAAC01000068; CAE8606.1; -; Genomic_DNA.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005788; Disulph. isom.
DR InterPro; IPR005792; Disulphide isom.
DR InterPro; IPR006662; Thioredo.
DR InterPro; IPR006663; Thioredox dom2.
DR InterPro; IPR012336; Thioredoxin-like.
PFam; PF00085; Thioredoxin.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01130; ER_PDI_fam; 1.
DR TIGRFAMs; TIGR01126; pdi_dom; 2.
DR PROSITE; PS00194; THIOREDOXIN; 2.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 55153 MW; 4358C2487E0A579B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
Db 277 QPKGKVL 283

RESULT 655
Q6PST0 ANCCA
ID Q6PST0 ANCCA PRELIMINARY; PRT; 493 AA.
AC Q6PST0;
DT 05-JUL-2004 (TREMREL. 27, Created)
DT 05-JUL-2004 (TREMREL. 27, Last sequence update)
DT 01-FEB-2005 (TREMREL. 29, Last annotation update)
DE Protein disulfide isomerase.
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG Hecke S.; Epe C.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY582130; AAS84455.1; -; mRNA.
DR EMBL; AY582129; AAS84454.1; -; Genomic_DNA.
DR HSSP; P00274; 1KEB.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.

Query Match 5.9%; Score 7; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
Db 277 QPKGKVL 283

RESULT 656
Q95PP0 OSTOS
ID Q95PP0 OSTOS PRELIMINARY; PRT; 493 AA.
AC Q95PP0;
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-MAR-2004 (TREMREL. 26, Last annotation update)
DE Disulfide isomerase.
GN Name=PD12;
OS Ostertagia ostertagi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Ostertaginae; Ostertagia.
OX NCBI_TaxID=6317;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG Geldhof P.B.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419174; CAD11865.1; -; mRNA.
DR HSSP; P07237; 1MEK.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005788; Disulphide isom.
DR InterPro; IPR005792; Disulphide isom.
DR InterPro; IPR006662; Thioredo.
DR InterPro; IPR006663; Thioredox dom2.
DR InterPro; IPR012336; Thioredoxin-like.
DR InterPro; IPR012335; Thioredoxin_fold.
PFam; PF00085; Thioredoxin; 3.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01130; ER_PDI_fam; 1.
DR TIGRFAMs; TIGR01126; pdi_dom; 2.
DR PROSITE; PS00194; THIOREDOXIN; 2.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 55153 MW; 4358C2487E0A579B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
Db 277 QPKGKVL 283

RESULT 657
Q95PP0 OSTOS
ID Q95PP0 OSTOS PRELIMINARY; PRT; 493 AA.
AC Q95PP0;
DT 01-DEC-2001 (TREMREL. 19, Created)
DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
DT 01-MAR-2004 (TREMREL. 26, Last annotation update)
DE Disulfide isomerase.
GN Name=PD12;
OS Ostertagia ostertagi.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Trichostrongyloidea; Haemonchidae; Ostertaginae; Ostertagia.
OX NCBI_TaxID=6317;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG Geldhof P.B.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ419174; CAD11865.1; -; mRNA.
DR HSSP; P07237; 1MEK.
DR GO; GO:0005783; C:endoplasmic reticulum; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005788; Disulphide isom.
DR InterPro; IPR005792; Disulphide isom.
DR InterPro; IPR006662; Thioredo.
DR InterPro; IPR006663; Thioredox dom2.
DR InterPro; IPR012336; Thioredoxin-like.
DR InterPro; IPR012335; Thioredoxin_fold.
PFam; PF00085; Thioredoxin; 3.
DR PRINTS; PR00421; THIOREDOXIN.
DR TIGRFAMs; TIGR01130; ER_PDI_fam; 1.
DR TIGRFAMs; TIGR01126; pdi_dom; 2.
DR PROSITE; PS00194; THIOREDOXIN; 2.
KW Isomerase.
SQ SEQUENCE 493 AA; 55007 MW; F194C55BFBC4313A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKVL 87
Db 277 QPKGKVL 283

```

```
Db 277 QFKGKVL 283
|||||
RESULT 657
Q524Y7 MAGGR
ID Q524Y7 MAGGR PRELIMINARY; PRT; 494 AA.
AC Q524Y7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG03116.4;
OS Magnaporthe grisea 70-15.
OC Sordariomycetes; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Eukaryota; Fungi; Ascomycota; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David K., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Girkie A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Kanat A., Kamyseselis M., Karlsson E.,
RA Kalls C., Kieu A., Kiser P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokitsang T., Lokitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marbella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Moghee T., Meldrim J., Meneus L.,
RA Masirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Mozes J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okawo O., O'leary S., Omotoho B.,
RA O'neil K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Rameasmy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,
RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoultsang Y., Topham K.,
RA Towey S., Tsamla T., Teomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whitaker C., Wilkinson J., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembek L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Magnaporthe grisea."
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;
RA Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
[3]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Zhu H., Blackmon B.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
```

```
EMBL: AACU01000663; EAA51521.1; -; Genomic_DNA.
DR InterPro; IPR003613; Ubox.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 3.
DR SMART; SM00504; UBox; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PSS0082; WD_REPEATS_2; 1.
DR PROSITE; PSS0294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; Ubl conjugation pathway; WD repeat.
SQ SEQUENCE 494 AA; 51769 MW; 233B9BEAC9AAB31E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 409 GVLAALA 415

RESULT 658
Q5YYI3 NOCPA
ID Q5YYI3 NOCPA PRELIMINARY; PRT; 499 AA.
AC Q5YYI3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-
DE alanyl-D-alanine ligase.
GN Name=murf; OrderedLocusNames=nfal7620;
OS Nocardia farcinica
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
RN NCBI_TaxID=37329;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IPM 10152;
RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IPM 10152."
RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930(2004).
DR EMBL; AP006618; BAD56608.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0008766; F:UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-d. .; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR InterPro; IPR005863; MurP.
DR InterPro; IPR004101; Mur_ligase_C.
DR InterPro; IPR000713; Mur_ligase_N.
DR Pfam; PF01225; Mur_ligase_1.
DR PF02875; Mur_ligase_C; 1.
DR TIGRFAMs; TIGR01143; murP; 1.
KW Complete proteome; Ligase.
SQ SEQUENCE 499 AA; 51063 MW; F5285AD149BB7CB0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 107 VLAALAA 113

RESULT 659
Q6NSK9 RHOPA
ID Q6NSK9 RHOPA PRELIMINARY; PRT; 499 AA.
AC Q6NSK9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
```


DE Hypothetical protein YjeF.
GN Name=YjeF; OrderedLocusNames=RPA2965;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R., Peres C.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572602; CAE28406.1; -; Genomic_DNA.
DR InterPro; IPR000631; UPF0031.
DR InterPro; IPR004443; YjeF_Nterm.
DR Pfam; PF01256; Carb_kinase; 1.
DR Pfam; PF03853; YjeF_N; 1.
DR TIGRFAMs; TIGR00196; YjeF_cterm; 1.
DR TIGRFAMs; TIGR00197; YjeF_aterm; 1.
DR PROSITE; PS01050; UPF0031_2; UNKNOWN_1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 499 AA; 51441 MW; 462738F0199F475E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 325 VLAALAA 331

RESULT 660
Q74F16 GEOSL
ID Q74F16 GEOSL PRELIMINARY; PRT; 499 AA.
AC Q74F16
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Membrane protein, putative.
GN OrderedLocusNames=GSU0622;
OS Geobacter sulfurreducens.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfuromonadales;
OC Geobacteraceae; Geobacter.
OX NCBI_TaxID=35554;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PCA / ATCC 51573;
RX PubMed=14671304; DOI=10.1126/science.1088727;
RA Methe B.A., Nelson K.E., Eisen J.A., Paulsen I.T., Nelson W.C.,
RA Heidelberg J.F., Wu D., Wu M., Ward M.L., Beanan M.J., Dodson R.J.,
RA Madupu R., Brinkac L.M., Daugherty S.C., DeBoy R.T., Durkin A.S.,
RA Gwinn M.L., Kolonay J.F., Sullivan J.A., Haft D.H., Selengut J.,
RA Daviden T.M., Zafar N., White O., Tran B., Romero C., Forberger H.A.,
RA Weidman J.F., Khouri H.M., Feldblyum T.V., Utterback T.R.,
RA Van Aken S.E., Lovley D.R., Fraser C.M.;
RT "Genome of Geobacter sulfurreducens: metal reduction in subsurface
RT environments.";
RL Science 302:1967-1969(2003).
DR EMBL; AE017180; AAR33953.1; -; Genomic_DNA.
DR TIGR; GSU0622; -.
DR InterPro; IPR001440; TPR.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Complete proteome.
SQ SEQUENCE 499 AA; 55785 MW; B4687A509B9EF0DC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAAY 27
DB 165 LAALAAY 171

RESULT 661
Q8KY30 STRCU
ID Q8KY30 STRCU PRELIMINARY; PRT; 500 AA.
AC Q8KY30
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative transporter.
OS Streptomyces collinus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=42684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM2012;
RA Saito H., Bruenker P., Martin R., Minas W.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF293355; AAM97376.1; -; Genomic_DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006857; P:oligopeptide transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005279; PspH_sympoort.
DR InterPro; IPR000109; PTR2.
DR Pfam; PF00854; PTR2; 1.
DR TIGRFAMs; TIGR00924; Yjdl_subl_fam; 1.
DR PROSITE; PS01022; PTR2_1; 1.
KW Transmembrane; Transport.
SQ SEQUENCE 500 AA; 51971 MW; 04651859AE867D41 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 500;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVL 21
DB 102 VLLGGVL 108

RESULT 662
Q8YPN6 ANASP
ID Q8YPN6 ANASP PRELIMINARY; PRT; 500 AA.
AC Q8YPN6
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE NADH dehydrogenase subunit 4.
GN Name=ndhd; OrderedLocusNames=alr4157;
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -!- FUNCTION: NDH-1 shuttles electrons from NAD(P)H, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is

believed to be plastoquinone. Couples the redox reaction to proton translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient (By similarity).
 -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) + plastoquinol.

EMBL; BA000019; BAB75856.1; -, Genomic_DNA.
 DR PIR; AF2325; AF2325.
 DR GO; GO:0008137; P:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0042773; P:ATP synthase coupled electron transport; IEA.
 DR InterPro; IPR003918; NADH_oxred4.
 DR InterPro; IPR010227; NDH I M.
 DR InterPro; IPR001750; Oxidored q1.
 DR Pfam; PF00361; Oxidored q1; 1.
 DR PRINTS; PR01437; NUOXDRDASE4.
 DR TIGRFAMs; TIGR01972; NDH I M; 1.
 KW Complete proteome; NAD; NADP; Oxidoreductase; Plastoquinone.
 SQ SEQUENCE 500 AA; 54130 MW; 9D2290E9B13212A0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 500;
 Best Local Similarity 100.0%; Pred. No. 5.9e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLAGVLA 22
 |||||
 Db 243 LLAGVLA 249

RESULT 663

Q7NLW5 GLOVI
 ID 07NLW5 GLOVI PRELIMINARY; PRT; 504 AA.
 AC Q7NLW5;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DE NADH dehydrogenase subunit 4.
 GN Name=ndhb; OrderedLocusNames=glri1004;
 OS Gloeobacter violaceus.
 OC Bacteria; Cyanobacteria; Gloeobacteria; Gloeobacteriales; Gloeobacter.
 OX NCBI_TaxID=33072;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=PCC 7421;
 RX MEDLINE=22977040; PubMed=14621292;
 RA Nakamura Y., Kaneo T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
 RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
 RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimo S.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a cyanobacterium that lacks thylakoids.";
 RL DNA Res. 10:137-145(2003).
 CC -1- FUNCTION: NDH-1 shuttles electrons from NAD(P)H, via FMN and iron-sulfur (Fe-S) centers, to quinones in the respiratory chain. The immediate electron acceptor for the enzyme in this species is believed to be plastoquinone. Couples the redox reaction to proton translocation (for every two electrons transferred, four hydrogen ions are translocated across the cytoplasmic membrane), and thus conserves the redox energy in a proton gradient (By similarity).
 CC -1- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) + plastoquinol.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
 EMBL; BA000045; BAC88945.1; -, Genomic_DNA.
 DR GO; GO:0008137; P:NADH dehydrogenase (ubiquinone) activity; IEA.
 DR GO; GO:0042773; P:ATP synthase coupled electron transport; IEA.
 DR InterPro; IPR003918; NADH_oxred4.
 DR InterPro; IPR010227; NDH I M.
 DR InterPro; IPR001750; Oxidored q1.
 DR Pfam; PF00361; Oxidored q1; 1.
 DR PRINTS; PR01437; NUOXDRDASE4.
 DR TIGRFAMs; TIGR01972; NDH I M; 1.
 KW Complete proteome; NAD; NADP; Oxidoreductase; Plastoquinone.
 SQ SEQUENCE 504 AA; 53482 MW; D365ACA1E285B831 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 504;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLAGVLA 22
 |||||
 Db 242 LLAGVLA 248

RESULT 664

Q4NLE6 9MICC
 ID Q4NLE6 9MICC PRELIMINARY; PRT; 505 AA.
 AC Q4NLE6;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Zinc-containing alcohol dehydrogenase superfamily.
 GN ORFNames=ArthDRAFT_4294;
 OS Arthrobacter sp. PE24.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Micrococcales; Micrococaceae; Arthrobacter.
 OX NCBI_TaxID=290399;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PE24;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
 RA Hammon N., Israni S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Arthrobacter sp. PB24.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PE24;
 RG US DOE Joint Genome Institute (JGI-PGF);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Arthrobacter sp. PB24.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.

CC -1- COPACTOR: Zinc (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 EMBL; AARG01000001; EAL98426.1; -, Genomic_DNA.
 DR InterPro; IPR002085; ADH_SF_Zn.
 DR InterPro; IPR002328; ADH_Zn.
 DR Pfam; PF00107; ADH zinc N; 1.
 DR PROSITE; PS00059; ADH_ZINC; UNKNOWN 1.
 KW Metal-binding; Oxidoreductase; Zinc.
 SQ SEQUENCE 505 AA; 52443 MW; 07900C44A4B66168 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 505;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLAGVLA 22
 |||||
 Db 392 LLAGVLA 398

RESULT 665

Q9Y9C5 AERPE
 ID Q9Y9C5 AERPE PRELIMINARY; PRT; 510 AA.
 AC Q9Y9C5;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein APE2362.
 GN OrderedLocusNames=APE2362;
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
 OC Desulfurococaceae; Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K3;
RM MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hogoyma A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RL crenarchaeon, Aeropyrum pernix K1.";
DR ENBL; BA000002; BAA81375.1; -; Genomic_DNA.
DR PIR; G72464; G72464.
DR InterPro; IPR000631; UPF0031.
DR InterPro; IPR004443; YjeF_Nterm.
DR Pfam; PF01256; Carb_kinase_1.
DR Pfam; PF03853; YjeF_N_1.
DR TIGRFAMs; TIGR00196; YjeF_cter; 1.
DR TIGRFAMs; TIGR00197; YjeF_Nterm; 1.
DR PROSITE; PS01050; UPF0031_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 510 AA; 52331 MW; FCBC34ACD1630281 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAAAY 27
|||||
DB 280 LAALAAAY 286

RESULT 666
Q9A748 CAUCR
ID Q9A748 CAUCR PRELIMINARY; PRT; 510 AA.
AC Q9A748;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein CC1878.
GN OrderedLocustNames=CC1878;
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
NCBI_TaxID=155892;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 19089 / CB15;
RM MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A.S., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson K.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
RA Emswolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
RA Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
DR ENBL; AE005861; AAK23853.1; -; Genomic_DNA.
DR PIR; A87482; A87482.
DR TIGR; CC1878; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001915; Peptidase_M48.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR011990; TPR-like_helical.
DR Pfam; PF01435; Peptidase_M48; 1.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Complete proteome; Hydrolase; Hypothetical protein; Metalloprotease;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein F19K16.18.
 GN Name=F19K16.18;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lin X., Kaul S., Town C.D., Benito M., Cressy T.H., Haas B.J., Wu D.,
 RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Uterback T.R., Fraser C.M.;
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Town C.D., Kaul S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC011717; AAG52233.1; -; Genomic_DNA.
 DR F1R; F96829; F96829.
 DR InterPro; IPR005512; DUF315.
 DR Pfam; PF03759; DUF315; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 515 AA; 58045 MW; DCD2755685CAA48A CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 515;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 20 VLAALAA 26
 Db 75 VLAALAA 81
 |||||
 |||||

 RESULT 669
 QAJTV7_CORJK PRELIMINARY; PRT; 519 AA.
 ID QAJTV7_CORJK PRELIMINARY;
 AC QAJTV7;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Putative sulfate permease family protein.
 GN ORFNames=jk1577;
 OS Corynebacterium jeikeium (strain K411).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OC NCBI_TaxID=306537;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RX PubMed=15968079; DOI=10.1128/JB.187.13.4671-4682.2005;
 RA Tauch A., Kaiser O., Hain T., Gessmann A., Weisshaar B.,
 RA Albermeier A., Bekel T., Bischoff N., Brune I., Chakraborty T.,
 RA Kalinowski J., Meyer F., Rupp O., Schneiker S., Viehoveer P.,
 RA Puhler A.;
 RT "Complete Genome Sequence and Analysis of the Multiresistant
 RT Nosocomial Pathogen Corynebacterium jeikeium K411, a Lipid-Requiring
 RT Bacterium of the Human Skin Flora.";
 RL J. Bacteriol. 187:4671-4682 (2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=K411;
 RA Linke B., Tauch A.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR931997; CA137750.1; -; Genomic_DNA.
 SQ SEQUENCE 519 AA; 54903 MW; F3D6FB05826D5F51 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 519;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
 Db 118 VLLGGVL 124
 |||||
 |||||

 RESULT 670
 Q6F2A2_MESFL PRELIMINARY; PRT; 521 AA.
 ID Q6F2A2_MESFL PRELIMINARY;
 AC Q6F2A2;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Arginine-ornithine APC transporter.
 GN OrderedLocuNames=Mf1015;
 OS Mesoplasma florum (Acholeplasma florum).
 OC Bacteria; Firmicutes; Mollicutes; Entomoplasmatales;
 OC Entomoplasmataceae; Mesoplasma.
 OC NCBI_TaxID=2151;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=L1 / ATCC 33453;
 RA Birren B.W., Stange-Thomann N., Hafez N., DeCaprio D., Fisher S.,
 RA Butler J., Elkins T., Kodira C.D., Major J., Wang S., Nicol R.,
 RA Nuebaum C.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017263; AA75371.1; -; Genomic_DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015556; F:C4-dicarboxylate transporter activity; IEA.
 DR GO; GO:0015740; P:C4-dicarboxylate transport; IEA.
 DR InterPro; IPR004669; DcuC.
 DR Pfam; PF03606; DcuC; 1.
 KW Complete proteome.
 SQ SEQUENCE 521 AA; 56421 MW; 934BE050A06F3C72 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 521;
 Best Local Similarity 100.0%; Pred. No. 6.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 16 LLLGGVLA 22
 Db 440 LLLGGVLA 446
 |||||
 |||||

 RESULT 671
 Q5YW97_NOCFA PRELIMINARY; PRT; 525 AA.
 ID Q5YW97_NOCFA PRELIMINARY;
 AC Q5YW97;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative oxidoreductase.
 GN OrderedLocuNames=nfa26970;
 OS Nocardia farcinica.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Nocardiaceae; Nocardia.
 OC NCBI_TaxID=37323;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=IFM 10152;
 RX PubMed=15466710; DOI=10.1073/pnas.0406410101;
 RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
 RA Shiba T., Hattori M.;
 RT "The complete genomic sequence of Nocardia farcinica IFM 10152.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:14925-14930 (2004).
 DR EMBL; AP006618; BAD57544.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0045285; C:ubiquinol-cytochrome-c reductase complex; IEA.
 DR GO; GO:0016491; F:ubiquinol-cytochrome-c reductase activity; IEA.
 DR GO; GO:0008121; F:ubiquinol-cytochrome-c reductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR006076; Fad_oxred.
 DR InterPro; IPR00205; NAD_BS.

```
DR InterPro; IPR005805; Rieseke.
DR InterPro; IPR005806; Rieseke reg.
DR InterPro; IPR003042; Rng_mnxygenase.
DR Pfam; PF01266; DAO; 1.
DR Pfam; PF00355; Rieseke; 1.
DR PRINTS; PR00162; RIESKE
DR PRINTS; PR00420; RINGMNOXGNASE.
KW Complete proteome.
SQ SEQUENCE 525 AA; 55485 MW; 05812774ABD5B5F8 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 191 VLAALAA 197

RESULT 672
Q8PFU5_XANAC PRELIMINARY; PRT; 526 AA.
AC Q8PFU5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Disulphide-isomerase.
GN OrderedLocusNames=XAC3878;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Faran C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral L.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR HSSP; A6012037; AAM38720.1; -; Genomic_DNA.
DR HSSP; O77404; 1073.
DR GO; GO:0005489; P:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006662; ThioRed.
DR InterPro; IPR006663; ThioRedox_dom2.
DR PRINTS; PR00421; THIOREDOXIN.
KW Complete proteome; Isomerase.
SQ SEQUENCE 526 AA; 56440 MW; 320201C56AE33A02 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 526;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 18 VLAALAA 24
```

```
RESULT 673
Q4H578_9DEIO PRELIMINARY; PRT; 529 AA.
AC Q4H578;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE ABC transporter precursor.
GN ORNames=DgeODRAFT_2520;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hamon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000029; EAL81443.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL 1 20 Potential.
SQ SEQUENCE 529 AA; 55726 MW; 4DF33887D7F34162 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 8 GVLAALA 14

RESULT 674
Q9KP41_VIBCH PRELIMINARY; PRT; 530 AA.
AC Q9KP41;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Thiamine ABC transporter, permease protein, putative.
GN OrderedLocusNames=VC2538;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.L.,
RA Ermolaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
```

```

RL Nature 406:477-483(2000).
CC -|- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL: AE004323; AAF95679.1; -; Genomic_DNA.
DR F1R: D82063; D82063.
DR TIGR: VC2536; -.
DR GO: GO:0019866; C:inner membrane; IEA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000515; BPD_transp.
DR InterPro: IPR005947; ThiP_ABC_transp.
DR Pfam: PF00528; BPD_transp_1; 2.
DR TIGRFAMs: TIGR01253; ThiP_1.
DR PROSITE: PS00928; ABC_TM1; 2.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 530 AA; 59046 MW; 9C3E3B4A23A55784 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 530;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 19 VLAALAA 25

RESULT 675
Q6SQV0_MANSM
ID Q6SQV0_MANSM PRELIMINARY; PRT; 533 AA.
AC Q6SQV0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DR OppA protein.
GN Name=oppA; OrderedLocusNames=MS2053;
OS Mannheimia succiniciproducens (strain MBEL55E).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Mannheimia.
OC NCBI_TaxID=221988;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15378067; DOI=10.1038/nbt1010;
RA Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
RA Kim C.H., Jeong H., Hur C.G., Kim J.J.;
RT "The genome sequence of the capnophilic rumen bacterium Mannheimia
RL succiniciproducens."
RL Nat. Biotechnol. 22:1275-1281(2004).
DR EMBL: AB016827; AAU38660.1; -; Genomic_DNA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
DR PROSITE: PS01040; SBP_BACTERIAL_5; 1.
KW Complete proteome.
SQ SEQUENCE 533 AA; 59530 MW; 70F788180BB850B4 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 533;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 QAOVIAH 80
Db 484 QAOVIAH 490

RESULT 676
TNSE_ECOLI
ID TNSE_ECOLI STANDARD; PRT; 538 AA.

```

```

AC P05845;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transposon Tn7 transposition protein tnsE (Protein D).
GN Name=tnsE;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90192166; PubMed=2156235;
RA Flores C., Qadri M.I., Lichtenstein C.;
RT "DNA sequence analysis of five genes; tnsA, B, C, D and E, required
RT for Tn7 transposition."
RL Nucleic Acids Res. 18:901-911(1990).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=87040763; PubMed=3022239;
RA Smith G.W., Jones P.;
RT "Tn7 transposition: a multigene process. Identification of a
RT regulatory gene product."
RL Nucleic Acids Res. 14:7915-7927(1986).
CC -|- FUNCTION: TnsABC + tnsD promote high-frequency insertion of Tn7
CC into a specific target site known as att-Tn7 whereas tnsABC + tnsE
CC promote low-frequency insertion into many different sites.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL: X17693; CAA35687.1; -; Genomic DNA.
DR EMBL: X04534; CAB56509.1; -; Genomic_DNA.
DR F1R: A25543; Q8ECD7.
KW DNA recombination; DNA-binding; Transposable element; Transposition.
FT DNA BIND 311 330 H-T-H motif (By similarity).
SQ SEQUENCE 538 AA; 61212 MW; DE34A3F141A1885B CRC64;

Query Match 5.9%; Score 7; DB 1; Length 538;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
Db 363 LGGVLA 369

RESULT 677
Q7AJH8_ECOLI
ID Q7AJH8_ECOLI PRELIMINARY; PRT; 538 AA.
AC Q7AJH8;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Transposition regulatory protein.
GN Name=tnsE;
OS Escherichia coli.
OS Plasmid R721.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN-K-12;
RX MEDLINE=93015772; PubMed=1400257;
RA Kim S., Komano T.;
RT "Nucleotide sequence of the R721 shufflon."
RL J. Bacteriol. 174:7053-7058(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.

```

```

RC STRAIN-K-12;
RA Sampa G., Motomura K., Masuda S., Yamaguchi T., Ando K., Oishi T.,
RA Furuya N., Komano T., Mizobuchi K.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002527; BAB12609.1; -; Genomic_DNA.
KW Plasmid.
SQ SEQUENCE 538 AA; 61211 MW; DE34A3F141A1885B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 538;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 23
Db 363 LGGVLA 369

RESULT 678
Q5QTU2_IDILO
ID Q5QTU2_IDILO PRELIMINARY; PRT; 543 AA.
AC Q5QTU2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE NAD synthase.
GN OrderedLocNames=I11186;
OS Idiomarina loihiensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Idiomarinaaceae; Idiomarina.
OC NCBI_TaxID=135577;
RN [1]

NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RP STRAIN=L2-TR / DSM 15497 / ATCC BAA-735;
RC PubMed=15596722; DOI=10.1073/pnas.0407638102;
RX Hou S., Saw J.H., Lee K.S., Freitas T.A., Belisle C., Kawarabayasi Y.,
RA Donachie S.P., Pikina A., Galperin M.Y., Koonin E.V., Makarova K.S.,
RA Omeichenko M.V., Sorokin A., Wolf Y.I., Li Q.X., Keum Y.S.,
RA Campbell S., Denery J., Aizawa S.-I., Shibata S., Malahoff A.,
RA Alam M.;
RA "Genome sequence of the deep-sea gamma-proteobacterium Idiomarina
RT loihiensis reveals amino acid fermentation as a source of carbon and
RT energy.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:18036-18041 (2004).
DR EMBL; AB017340; AAV62026.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016810; F:hydrolase activity, acting on carbon-nitrog. .; IEA.
DR GO; GO:0003952; F:NAD+ synthase (glutamine-hydrolyzing) activity; IEA.
DR GO; GO:0009435; F:NAD biosynthesis; IEA.
DR GO; GO:0006807; P:nitrogen compound metabolism; IEA.
DR InterPro; IPR003694; NAD_synthase.
DR InterPro; IPR003010; N1lse/CNhydtaa.
DR Pfam; PF00795; CN hydrolase; 1.
DR Pfam; PF02540; NAD_synthase; 1.
DR TIGRFAMs; TIGR00552; nadE; 1.
DR PROSITE; PS0263; CN HYDROLASE; 1.
KW Complete proteome.
SQ SEQUENCE 543 AA; 60407 MW; E51EPDBEF88947D CRC64;

Query Match 5.9%; Score 7; DB 2; Length 543;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CWVIVGH 39
Db 78 CWVIVGH 84

RESULT 679
Q92NE8_RHIME
ID Q92NE8_RHIME PRELIMINARY; PRT; 543 AA.
AC Q92NE8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE PUTATIVE ABC TRANSPORTER ATP-BINDING PROTEIN.
GN OrderedLocNames=R02257; ORFNames=SMC01645;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OC NCBI_TaxID=382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1021;
EX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Gohrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
CC -!- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC -!- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL; AL591790; CAC46836.1; -; Genomic_DNA.
DR HSSP; Q58206; 1L2T
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER 1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER 2; 2.
KW ATP-binding; Complete proteome; Inner membrane; Membrane;
KW Nucleotide-binding; Transport.
SQ SEQUENCE 543 AA; 59387 MW; B22E44B7A62E55 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 543;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLGL 89
Db 38 KGKVLGL 44

RESULT 680
Q70FGI_PSEPU
ID Q70FGI_PSEPU PRELIMINARY; PRT; 546 AA.
AC Q70FGI;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative mobA protein.
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OC NCBI_TaxID=303;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Carl B., Arnold A., Hauer B., Fetzner S.;
RT "Sequence and transcriptional analysis of a gene cluster of
RT Pseudomonas putida 86 involved in quinoline degradation.";
RL Gene 331:177-188 (2004).
DR EMBL; AJ583091; CAS47360.1; -; Genomic_DNA.
DR GO; GO:0006777; P:Mo-molybdopterin cofactor biosynthesis; IEA.
DR InterPro; IPR001453; MOCF_bios.
DR InterPro; IPR012184; MPTB_MobAlake.
DR PIRSF; PIRSF036626; MPTB_MobAlake; 1.
SQ SEQUENCE 546 AA; 57475 MW; 2C4B489C391391A0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 546;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;

```



```

RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL Mesorhizobium loti."
CC DNA Res. 7:331-338(2000).
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (By similarity).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR EMBL: BA000012; BAB51625.1; -; Genomic_DNA.
DR HSPSP; Q58206; 1127.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0016887; F:ATPase activity; IEA.
DR GO: GO:0000156; F:nucleotide binding; IEA.
DR InterPro: IPR003593; AAA ATPase.
DR InterPro: IPR003439; ABC_transp_like.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transporter; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE: PS00893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Complete proteome; Inner membrane; Membrane;
KW Nucleotide-binding; Transport.
SQ SEQUENCE 551 AA; 59808 MW; 22D36ABE2F659EEF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGRVLGL 89
Db 45 KGRVLGL 51

RESULT 684
Q4UVQ9 XANCP PRELIMINARY; PRT; 552 AA.
AC Q4UVQ9.
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Chemotaxis protein.
OS ORFNames=XC1801.
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OC NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: CP000050; AAY4864.1; -; Genomic_DNA.
SQ SEQUENCE 552 AA; 58401 MW; E6D59666D54B5436 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 552;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 202 VLLGGVL 208

RESULT 685
Q8P8C8 XANCP PRELIMINARY; PRT; 552 AA.
AC Q8P8C8.
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Chemotaxis protein.

```

```

GN Name=mcp;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Canavan F., Cardozo J., Chambergo F.,
RA Clapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorzy H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL: AB012339; AAM41593.1; -; Genomic_DNA.
DR HSPSP; P02942; 1QV7.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0004871; F:signal transducer activity; IEA.
DR GO: GO:0006935; P:chemotaxis; IEA.
DR GO: GO:0007165; P:signal transduction; IEA.
DR InterPro: IPR004089; Chmtaxis trans.
DR InterPro: IPR003660; His_kin_HAMP.
DR InterPro: IPR004090; Me_Chemotaxis.
DR Pfam: PF00672; HAMP; 1.
DR Pfam: PF00015; MCPsignal; 1.
DR PRINTS: PR00260; CHEMTRNSDUCR.
DR ORFNames=XC1801.
DR SMART: SM00304; HAMP; 1.
DR SMART: SM00283; MA; 1.
DR PROSITE: PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE: PS00885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 552 AA; 58401 MW; E6D59666D54B5436 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 552;
Best Local Similarity 100.0%; Pred. No. 6.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
Db 202 VLLGGVL 208

RESULT 686
Q6LIA9 PHOPR PRELIMINARY; PRT; 558 AA.
AC Q6LIA9.
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative electron transfer flavoprotein-ubiquinone
DE oxidoreductase.
GN Name=RSCL567; OrderedLocusNames=PPPR1099;
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX PubMed=15746425; DOI=10.1126/science.1103341;
RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N.,

```

```

RA Lauro F.M., Cestaro A., Malacrida G., Simonati B., Cannata N.,
RA Romualdi C., Bartlett D.H., Valle G.;
RT "Life at depth: Photobacterium profundum genome sequence and
RT expression analysis.";
RL Science 307:1459-1461(2005).
DR EMBL; CR378678; CAG22971.1; -; Genomic DNA.
DR GO; GO:0004174; P:electron-transferring-flavoprotein dehydrog. . .; IEA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR007859; EFTD.
DR InterPro; IPR006076; Pad_oxred.
DR InterPro; IPR003042; Rng_mnoxygenase.
DR Pfam; PF01266; DAO; 1.
DR Pfam; PF05187; ETP_QO; 1.
DR PRINTS; PR00420; RINGNOXGNASE.
DR Complete proteome; Ubiquinone.
SQ SEQUENCE 558 AA; 60998 MW; 45D561B7CA349B08 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 558;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLAGVLA 22
| | | | |
Db 404 LLAGVLA 410

RESULT 687
Q9HRT5_HALSA
ID Q9HRT5_HALSA PRELIMINARY; PRT; 559 AA.
AC Q9HRT5
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Vng0553C.
GN OrderedLocNames=VNG0553C;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
[1]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Fohlsechroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005006; AAG19073.1; -; Genomic_DNA.
DR FIK; E84213; E84213.
KW Complete proteome.
SQ SEQUENCE 559 AA; 55157 MW; C89F0D8D3AF59EB5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 559;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 111 VLAALAA 117

RESULT 688
Q8L800_FLABI
ID Q8L800_FLABI PRELIMINARY; PRT; 563 AA.
AC Q8L800;

```

```

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glutamate/malate translocator.
GN Name=Dit2;
OS Flaveria bidentis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Tageteae; Flaveria.
OX NCBI_TaxID=4224;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Rowbotham-Bradford;
RX MEDLINE=22770869; PubMed=12887583;
RX DOI=10.1046/j.1365-3113.2003.01806.x;
RA Renne P., Dreesen U., Hebbeker U., Hille D., Plugge U.I., Westhoff P.,
RA Weber A.P.;
RT "The Arabidopsis mutant dct is deficient in the plastidic
RL Plant J. 35:316-331(2003).
DR EMBL; AY123845; AM89395.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_sympor.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR TIGRFAMs; TIGR00785; dase; 1.
DR SEQUENCE 563 AA; 60327 MW; CACE4A9D164CAB05 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 563;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
| | | | |
Db 495 GVLAALA 501

RESULT 689
Q5UFU3_MIMIV
ID Q5UFU3_MIMIV PRELIMINARY; PRT; 563 AA.
AC Q5UFU3;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MIMI_L116;
OS Mimivirus.
OC Viruses; dsDNA viruses, no RNA stage; Mimivirus.
OX NCBI_TaxID=212035;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Rowbotham-Bradford;
RX MEDLINE=22550848; PubMed=12663918; DOI=10.1126/science.1081867;
RA La Scola B., Audic S., Robert C., Jungang L., de Lamballerie X.,
RA Drancourt M., Birtles R., Claverie J.M., Raoult D.;
RT "A giant virus in amoebae.";
RL Science 299:2033-2033(2003).
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=Rowbotham-Bradford;
RX PubMed=15486256; DOI=10.1126/science.1101485;
RA Raoult D., Audic S., Robert C., Abergel C., Renesto P., Ogata H.,
RA La Scola B., Susan M., Claverie J.M.;
RT "The 1.2-Mb Genome Sequence of Mimivirus.";
RL Science 306:1344-1350(2004).
DR EMBL; AY653733; AAV50391.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 563 AA; 66608 MW; 7BB0901D813D67E9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 563;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 43 GCKPAIV 49
Db 516 GCKPAIV 522

RESULT 690
Q7Q9Y0 ANOGA
ID Q7Q9Y0 ANOGA PRELIMINARY; PRT; 564 AA.
AC Q7Q9Y0;
DT 01-MAR-2004 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE ENSANGP00000011688 (Fragment).
GN ORFNames=ENSANG000000009199;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008898; EAA09180.2; -; Genomic_DNA.
FT NON TER 564 564
SQ SEQUENCE 564 AA; 60188 MW; 87A0455F8E908925 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 564;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 388 VLAALAA 394

RESULT 691
Q8L7Z9 SPIOL
ID Q8L7Z9 SPIOL PRELIMINARY; PRT; 564 AA.
AC Q8L7Z9;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Glutamate/malate translocator.
GN Name=Dit1;
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Amaranthaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=22770869; PubMed=12887583;
RX DOI=10.1046/j.1365-3113X.2003.01806.x;
RA Renee P.; Dressen U.; Hebbeker U.; Hille D.; Flugge U.I.; Westhoff P.;
RA Weber A.P.;
RT "The Arabidopsis mutant dct is deficient in the plastidic
RT glutamate/malate translocator Dit2.";
RL Plant J. 35:316-331(2003).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Weber A.; Hebbeker U.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY123846; AAM89396.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na/sul_symp; IEA.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR TIGRFAMs; TIGR00785; dass; 1.
SQ SEQUENCE 564 AA; 59873 MW; 6DC64D3E7925F209 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 564;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
Db 496 GVLAALA 502

RESULT 692
Q7VZ95 BORPE
ID Q7VZ95 BORPE PRELIMINARY; PRT; 568 AA.
AC Q7VZ95;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Methyl-accepting chemotaxis protein I.
GN Name=ter; Synonyms=ched; OrderedLocusNames=BP1030;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J.; Sebahia M.; Preston A.; Murphy L.D.; Thomson N.R.;
RA Harris D.E.; Holden M.T.G.; Churcher C.M.; Bentley S.D.; Mungall K.L.;
RA Cardeno-Tarraga A.-M.; Temple L.; James K.D.; Harris B.; Quail M.A.;
RA Achtman M.; Atkin R.; Baker S.; Basham D.; Bason N.; Cherevach I.;
RA Chillingworth T.; Collins M.; Cronin A.; Davis P.; Doggett J.;
RA Feltwell T.; Gobie A.; Hamlin N.; Hauser H.; Holroyd S.; Jagels K.;
RA Leather S.; Moule S.; Norberczak H.; O'Neill S.; Ormond D.; Price C.;
RA Rabinowitsch E.; Rutter S.; Sanders M.; Saunders D.; Seeger K.;
RA Sharp S.; Simmonds M.; Skelton J.; Squares R.; Squares S.; Stevens K.;
RA Unwin L.; Whitehead S.; Barrell B.G.; Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640414; CAE41330.1; -; Genomic_DNA.
DR SMR; Q7VZ95; 309-532.
GO; GO:0019866; C:inner membrane; IEA.
GO; GO:0016021; C:integral to membrane; IEA.
GO; GO:0004871; F:signal transducer activity; IEA.
GO; GO:0006935; P:chemotaxis; IEA.
GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis trans.
DR InterPro; IPR003660; His kin_HAMP.
DR InterPro; IPR004090; Me_Chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR PRINTS; PR00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00283; HAM; 1.
DR SMART; SM00304; HAMP; 1.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS00885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 568 AA; 59815 MW; 0061CE25059FD4FC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 568;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 GGVLAAL 24

```

```
Db 214 GGVLAAL 220
|||||
RESULT 693
Q7MAA6 BORPA
ID Q7MAA6 BORPA PRELIMINARY; PRT; 568 AA.
AC Q7MAA6
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Methyl-accepting chemotaxis protein I.
GN Name=ter; Synonyms=chb; OrderedLocusNames=BPPI474;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RL EMBL; BX640427; CAE36776.1; -; Genomic_DNA.
DR SMR; Q7MAA6; 309-532.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis trans.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004090; Me_Chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00283; MA; 1.
DR SMART; SM00304; HAMP; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 568 AA; 359FF738C4588539 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 568;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 214 GGVLAAL 220
|||||
RESULT 694
Q7WJ9 BORBR
ID Q7WJ9 BORBR PRELIMINARY; PRT; 568 AA.
AC Q7WJ9
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Methyl-accepting chemotaxis protein I.
GN Name=ter; Synonyms=chb; OrderedLocusNames=BB2548;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RL EMBL; BX640427; CAE36776.1; -; Genomic_DNA.
DR SMR; Q7WJ9; 309-532.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis trans.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004090; Me_Chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00283; MA; 1.
DR SMART; SM00304; HAMP; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 568 AA; 359FF738C4588539 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 568;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 214 GGVLAAL 220
|||||
RESULT 694
Q7WJ9 BORBR
ID Q7WJ9 BORBR PRELIMINARY; PRT; 568 AA.
AC Q7WJ9
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Methyl-accepting chemotaxis protein I.
GN Name=ter; Synonyms=chb; OrderedLocusNames=BB2548;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RL EMBL; BX640444; CAE33042.1; -; Genomic_DNA.
DR SMR; Q7WJ9; 309-532.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis trans.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004090; Me_Chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00283; MA; 1.
DR SMART; SM00304; HAMP; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 568 AA; 59767 MW; CCC24418C145C9B8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 568;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 214 GGVLAAL 220
|||||
RESULT 695
Q6LSF6 ORYSA
ID Q6LSF6 ORYSA PRELIMINARY; PRT; 572 AA.
AC Q6LSF6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative auxin efflux carrier.
GN Name=OJ1126.B10.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Heu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
```

```
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RA "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RL EMBL; BX640444; CAE33042.1; -; Genomic_DNA.
DR SMR; Q7WJ9; 309-532.
DR GO; GO:0019866; C:inner membrane; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR004089; Chmtaxis trans.
DR InterPro; IPR003660; His_kin_HAMP.
DR InterPro; IPR004090; Me_Chemotaxis.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR PRINTS; PR00260; CHEMTRNSDUCR.
DR SMART; SM00283; MA; 1.
DR SMART; SM00304; HAMP; 1.
DR PROSITE; PS50111; CHEMOTAXIS_TRANSDUC_2; 1.
DR PROSITE; PS50885; HAMP; 1.
KW Complete proteome.
SQ SEQUENCE 568 AA; 59767 MW; CCC24418C145C9B8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 568;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 214 GGVLAAL 220
|||||
RESULT 695
Q6LSF6 ORYSA
ID Q6LSF6 ORYSA PRELIMINARY; PRT; 572 AA.
AC Q6LSF6
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative auxin efflux carrier.
GN Name=OJ1126.B10.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE
RC Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Heu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
```

RP NUCLEOTIDE SEQUENCE.
RA Chow T.-Y.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC098571; AAT39149.1; -, Genomic_DNA.
DR Gramene; Q6L5F6; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 572 AA; 60827 MW; F85CD5789A07B68F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 572;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 82 VLAALAA 88

RESULT 696
COX1B_STRCO
ID COX1B_STRCO STANDARD; PRT; 573 AA.
AC Q9K451;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Putative cytochrome c oxidase polypeptide I-beta (EC 1.9.3.1)
DE (Cytochrome aa3 subunit 1-beta).
GN Name=ctaD2; OrderedLocuNames=SC07234; ORFNames=C2H12.33;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1302;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.P., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RP "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- FUNCTION: Cytochrome c oxidase is the component of the respiratory
chain that catalyzes the reduction of oxygen to water. Subunits 1-
3 form the functional core of the enzyme complex. CO 1 is the
catalytic subunit of the enzyme. Electrons originating in
cytochrome c are transferred via the copper A center of subunit 2
and heme A of subunit 1 to the bimetallic center formed by heme A3
and copper B (By similarity).
CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferrocyclochrome
c + 2 H(2)O.
CC -1- COPACTOR: Binds 1 copper B per subunit (By similarity).
CC -1- COPACTOR: Binds 2 heme groups per subunit (By similarity).
CC -1- PATHWAY: Respiratory chain; terminal step.
CC -1- SUBUNIT: Associates with subunits II, III and IV to form
cytochrome c oxidase (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the heme-copper respiratory oxidase family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR EMBL; AL939130; CAB94657.1; -, Genomic_DNA.
DR HSP; P18401; 1PFT.
DR InterPro; IPR000883; COX1.
DR PANTHER; PTHR10422; COX1; 1.
DR Pfam; PF00115; COX1; 1.
DR PRINTS; PR01165; CYCOXIDASEI.
DR PROSITE; PS00855; COX1; 1.
DR PROSITE; PS00077; COX1_CUB; 1.
KW Complete proteome; Copper; Electron transport; Heme; Iron;
KW Metal-binding; Oxidoreductase; Respiratory chain; Transmembrane;
KW Transport.

FT TRANSMEM 53 73 Potential.
FT TRANSMEM 103 123 Potential.
FT TRANSMEM 141 161 Potential.
FT TRANSMEM 188 208 Potential.
FT TRANSMEM 227 247 Potential.
FT TRANSMEM 272 292 Potential.
FT TRANSMEM 304 324 Potential.
FT TRANSMEM 329 349 Potential.
FT TRANSMEM 373 393 Potential.
FT TRANSMEM 412 432 Potential.
FT TRANSMEM 447 467 Potential.
FT TRANSMEM 490 510 Potential.
FT METAL 100 100 Iron (heme A axial ligand) (By
similarity).
FT METAL 278 278 Copper B (By similarity).
FT METAL 282 282 Copper B (By similarity).
FT METAL 327 327 Copper B (By similarity).
FT METAL 328 328 Copper B (By similarity).
FT METAL 411 411 Iron (heme A3 axial ligand) (By
similarity).
FT METAL 413 413 Iron (heme A axial ligand) (By
similarity).
FT CROSSLNK 278 282 1'-histidyl-3'-tyrosine (His-Tyr) (By
similarity).
FT SQ SEQUENCE 573 AA; 64041 MW; 1FC5DC79FD6DA220 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 573;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
| | | | |
Db 67 LLGGVLA 73

RESULT 697
Q8X8A7_EC057
ID Q8X8A7_EC057 PRELIMINARY; PRT; 582 AA.
AC Q8X8A7; Q7A9A4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative frv operon regulatory protein.
GN Name=frvR; OrderedLocuNames=ECs4823; #5440;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Rose D.J., Hayek J., Klink S., Boutin A., Shao Y., Miller L.,
RA Posfai G., Mackett J., Davis N.W., Lim A., Dimalanta E.T., Potamoukis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RP "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.

```

RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHBC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT *Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12";
RL DNA Res. 8:11-22(2001);
DR EMBL; AF005174; AAC59090.1; -; Genomic_DNA.
DR EMBL; BA000007; BAB38246.1; -; Genomic_DNA.
DR PIR; F86078; F86078.
DR PIR; G91231; G91231.
DR GO; GO:0005351; P:sugar porter activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002178; PTS_EIIA_2.
DR Pfam; PF00359; PTS_EIIA_2; 1.
KW Complete proteome.
SQ SEQUENCE 582 AA; 65985 MW; 6088240EA76DD892 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGLLQR 92
Db 252 VLGLLQR 258

RESULT 698
Q8L7Z8_TOBAC
ID Q8L7Z8_TOBAC PRELIMINARY; PRT; 593 AA.
AC Q8L7Z8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glutamate/malate translocator.
DN Name=DIT2;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiales; Solanales; Solanaceae; Nicotiana.
OC NCBI_TaxID=4097;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22770869; PubMed=12887583;
RX DOI=10.1046/j.1365-3113.2003.01806.x;
RA Renne P., Dreesen U., Hebbeker U., Hille D., Flugge U.I., Westhoff P.,
RA Weber A.P.;
RT "The Arabidopsis mutant dct is deficient in the plastidic
RT glutamate/malate translocator DIT2.";
RL Plant J. 35:316-331(2003).
DR EMBL; AY123847; AAM89397.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001898; Na_sulph_symp; 1.
DR Pfam; PF00939; Na_sulph_symp; 1.
DR TIGRFAMs; TIGR00785; dase; 1.
SQ SEQUENCE 583 AA; 62138 MW; BA0B5C6988FD828F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 515 GVLAALA 521

RESULT 699
Q9KAF1_BACHD

```

```

ID Q9KAF1_BACHD PRELIMINARY; PRT; 585 AA.
AC Q9KAF1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ABC transporter (ATP-binding protein).
GN OrderedLocusNames=BH2336;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxID=86665;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT *Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL; BA000004; BAB06055.1; -; Genomic_DNA.
DR PIR; H83941; H83941.
DR HSP; P08716; 1MT0.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS00929; ABC_TMIF; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 585 AA; 65051 MW; 71EP4A8C56087C6C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 585;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 61 GVLAALA 67

RESULT 700
Q5ASQ5_EMENI
ID Q5ASQ5_EMENI PRELIMINARY; PRT; 588 AA.
AC Q5ASQ5;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein.
GN ORENAMES=AN8675.2;
OS Aspergillus nidulans FGSC A4.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; Emericella.
OC NCBI_TaxID=227321;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,

```

RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gadyana S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
 RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Sirovov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 RA Vasiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Genome Sequence of Aspergillus nidulans";
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC preliminary data.
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACD01000159; EAA60097.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 588 AA; 64393 MW; 5815CE2F24416A74 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 588;
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 VLLGGVL 21
 Db 285 VLLGGVL 291

RESULT 701
 Q9XBFP MYXXA
 ID Q9XBFP MYXXA PRELIMINARY; PRT; 589 AA.
 AC Q9XBFP;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Serine/threonine kinase PKN9.
 GN Name=PKN9;
 OS Myxococcus xanthus.
 OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
 OC Cytophactereae; Myxococcaceae; Myxococcus.
 OX NCBI_TaxID=34;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DZP1;
 RX MEDLINE=97197191; PubMed=9044280;
 RA Haulon W.A., Inouye M., Inouye S.;
 RT "Pkn9, a Ser/Thr protein kinase involved in the development of
 RT Myxococcus xanthus";
 RL Mol. Microbiol. 23:459-471 (1997).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=DZP1;
 RA Inouye S., Jain R., Ueki T., Nariya H., Xu C., Hsu M.,
 RA Munoz-Dorado J., Farez-Vidal E., Inouye M.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AF159694; AAD42859.1; -; Genomic_DNA.
 DR HSPB; P71584; 1MRU.
 DR Phosite; Q9XBFP3; -;
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.

DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
 DR PROSITE; PS00108; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.
 KW ATP-binding; Kinase; Nucleotide-binding;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 589 AA; 62731 MW; 3E80DEB636FAFEA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 589;
 Best Local Similarity 100.0%; Pred. No. 6.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 Db 176 VLAALAA 182

RESULT 702
 Q9VRN1 DROME
 ID Q9VRN1 DROME PRELIMINARY; PRT; 592 AA.
 AC Q9VRN1; Q8SYP6;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 10-MAY-2005 (TREMBLrel. 30, Last annotation update)
 DE CG10590-PA (RE4876P).
 GN ORFNames=CG10590;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bensch P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Bussan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zheng L.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426065; PubMed=12537568;
RA Celnikier S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celnikier S.E.,
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomic perspective.";
RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
[4]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Battencourt B.R., Celnikier S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.,
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
[5]
RN NUCLEOTIDE SEQUENCE.
RG Berkeley Drosophila Genome Project;
RA Celnikier S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
RA Hoskins R., Stapleton M., Pacleb J., Park S., Svirskas R., Smith E.,
RA Yu C., Rubin G.,
RT "Drosophila melanogaster release 4 sequence.";
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[6]
RN NUCLEOTIDE SEQUENCE.
RG FlyBase;
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
[7]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.,
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- INTERACTION:
CC QSVHP1:CG11753; NRExp1; IntAct=EBI-167716, EBI-194830;
CC EMBL; A003565; AAF50762.2; -; Genomic_DNA.
CC EMBL; AY071401; AAL49023.1; -; mRNA.
CC IntAct; QSVRN1; -;
CC Ensembl; CG10590; Drosophila melanogaster.
CC FlyBase; FBGN0035622; CG10590.
CC GO; GO:0016021; C:integral to membrane; IEA.
CC GO; GO:0005215; P:transporter activity; IEA.
CC GO; GO:0005215; P:transporter activity; IEA.
CC InterPro; IPR004240; EMP70.
CC PANTHER; PTHR10766; EMP70; 1.
CC Pfam; PF02990; EMP70; 1.
SQ SEQUENCE 592 AA; 67808 MW; 02935B062DBEC65B CRC64;
Query Match 5.9%; Score 7; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGVLA 21
Db 453 VLLGGVLA 459
RESULT 703
Q97IH2_CLOAB
ID Q97IH2_CLOAB PRELIMINARY; PRT; 592 AA.
AC Q97IH2;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Carbon starvation protein.
GN OrderedLocustNames=CAC1669;
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 824 / DSM 792 / VVM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
DOI=10.1128/JB.183.16.4823-4838.2001;
RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
RA Tatusov R.L., Sabathe P., Doucette-Stamm L.A., Soucaille P.,
RA Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.,
RT "Genome sequence and comparative analysis of the solvent-producing
RT bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007676; AAK79635.1; -; Genomic_DNA.
DR FJX; H97105; H97105.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009267; P:cellular response to starvation; IEA.
DR InterPro; IPR003706; CstA.
DR Pfam; PF02554; CstA; 1.
KW Complete proteome.
SQ SEQUENCE 592 AA; 64591 MW; 1C729F96A2462325 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 592;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLLGGVLA 22
Db 90 LLLGGVLA 96
RESULT 704
Q75UN9_HALDI
ID Q75UN9_HALDI PRELIMINARY; PRT; 594 AA.
AC Q75UN9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Cellulase.
GN Name=Hdcel;
OS Haliotis discus hannai.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Vetigastropoda; Haliotidae; Haliotis.
OX NCBI_TaxID=42344;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Suzuki K., Ojima T., Nishita K.,
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB125892; BAD01504.1; -; Genomic_DNA.
DR HSP; P37700; 1K72
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001701; Glyco_hydro_9.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
SQ SEQUENCE 594 AA; 64615 MW; 039AB5FD800A051 CRC64;


```
Query Match      5.9%; Score 7; DB 2; Length 594;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 5 VLAALAA 11

RESULT 705
Q86M37 HALDI
ID Q86M37_HALDI PRELIMINARY; PRT; 594 AA.
AC Q86M37;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cellulase
GN Name=hcdcl-1;
OS Haliotis discus (Abalone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Vetigastropoda; Haliotidae; Haliotidae; Haliotis.
OX NCBI_TaxID=36094;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12581217; DOI=10.1046/j.1432-1033.2003.03443.x;
RA Suzuki K., Ojima T., Nishita K.;
RT "Purification and cDNA cloning of a cellulase from abalone Haliotis
discus hannai.";
RL Euk. J. Biochem. 270:771-778 (2003).
DR EMBL; AB092978; BAC67186.1; -; mRNA.
DR HSSP; P26221; 1TF4.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . . ; IEA.
DR GO; GO:0030247; F:polysaccharide binding; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001291; Carb_cell_bind.
DR InterPro; IPR001701; Glyco_hydro_9.
DR Pfam; PF00759; Glyco_hydro_9; 1.
DR PROSITE; PS00698; GLYCOSYL_HYDROL_F9_2; 1.
SQ SEQUENCE 594 AA; 64601 MW; 59C7F3F1D9D61E1D CRC64;

Query Match      5.9%; Score 7; DB 2; Length 594;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 5 VLAALAA 11

RESULT 706
Q9RRP4 DEIRA
ID Q9RRP4_DEIRA PRELIMINARY; PRT; 603 AA.
AC Q9RRP4;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Nucleic acid-binding protein, putative, HRDC family.
GN OrderedLocNames=DR2444;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
```

```
Fraser C.M.;
"Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577 (1999).
DR EMBL; AE002074; AAF11987.1; -; Genomic_DNA.
DR PIR; H75272; H75272.
DR TIGR; DR2444; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR002121; HRDC.
DR Pfam; PF00570; HRDC; 1.
DR PROSITE; PS09667; HRDC; 1.
KW Complete proteome.
SQ SEQUENCE 603 AA; 64245 MW; 608C3DC18B11CD67 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 603;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 23
DB 227 LGGVLA 233

RESULT 707
P72607 SYN3
ID P72607 SYN3 PRELIMINARY; PRT; 605 AA.
AC P72607;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE ABC transporter.
GN OrderedLocNames=slr1488;
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasanoto S., Kimura T.,
RA Hiseouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimp S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
Synechocystis sp. strain PCC6803. II. Sequence determination of the
entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136 (1996).
DR EMBL; BA000022; BAAL6607.1; -; Genomic_DNA.
DR PIR; S74455; S74455.
DR HSSP; Q03518; 1JU7.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0042626; F:ATPase activity, coupled to transmembrane m. . . ; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR011527; ABC_membrane_1.
DR InterPro; IPR001140; ABC_TM_transp.
DR InterPro; IPR003439; ABC_transp_like.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR ProDom; PD000006; ABC_transporter; 1.
DR SMART; SM00382; AAA_1.
DR PROSITE; PS0929; ABC_TM1F; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
KW Complete proteome.
SQ SEQUENCE 605 AA; 67066 MW; AFD44380D8CFDA4A CRC64;

Query Match      5.9%; Score 7; DB 2; Length 605;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
```

```
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 84 GKVLGLL 90
Db 389 GKVLGLL 395

RESULT 708
Y4PA_RHISN STANDARD; PRT; 609 AA.
AC P5610;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Putative transcriptional regulatory protein y4pa.
GN ORFNames=y4pa;
OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234a.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Parret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -!- FUNCTION: Probable transcriptional regulator that acts in
CC conjugation with sigma-54.
CC -!- SIMILARITY: Contains 1 sigma-54 factor interaction domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AE000089; AAB91811.1; -; Genomic_DNA.
DR InterPro; IPR003593; AAA_Arpase.
DR InterPro; IPR002197; HTH_Fis.
DR InterPro; IPR002078; Sig54_interact.
DR Pfam; PF02954; HTH_8; 1.
DR Pfam; PF00158; Sigma54_activat; 1.
DR PRINTS; PR01590; HTHFIS.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR01199; HTH_fis; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; FALSE NEG.
DR PROSITE; PS00676; SIGMA54_INTERACT_2; FALSE NEG.
DR PROSITE; PS00688; SIGMA54_INTERACT_3; FALSE NEG.
DR PROSITE; PS50045; SIGMA54_INTERACT_4; 1.
KW ATP-binding; DNA-binding; Hypothetical protein; Nucleotide-binding;
KW Plasmid; Sensory transduction; Transcription;
KW Transcription regulation; Two-component regulatory system.
FT DOMAIN 313 533 Sigma-54 factor interaction (Potential).
FT NP BIND 395 404 ATP (Potential).
FT DNA_BIND 578 597 H-T-H motif (By similarity).
SQ SEQUENCE 609 AA; 66700 MW; 8CC727E67D508F36 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 609;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 85 KVLGLLQ 91
Db 387 KVLGLLQ 393

RESULT 709
Q41ZJ5_AZOVI
ID Q41ZJ5_AZOVI PRELIMINARY; PRT; 614 AA.
AC Q41ZJ5;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
us-09-638-693a-36_copy_16_133.olig.rup
```

```
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Type II and III secretion system protein precursor.
GN ORFNames=AvindRAFT_4287;
OS Azotobacter vinelandii AVOP.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Azotobacter.
OX NCBI_TaxID=322710;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hamon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Azotobacter vinelandii
RT AVOP.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Azotobacter vinelandii
RT AVOP.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AVOP;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Dettler C., Glavina T.,
RA Hamon N., Israni S., Pitluck S., Richardson P.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABU03000003; EAM06112.1; -; Genomic_DNA.
DR Signal.
KW SIGNAL 1 29 Potential.
FT SIGNAL 614 614 Potential.
SQ SEQUENCE 614 AA; 66835 MW; 18A3114400D2153C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 614;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 15 VLLGGVL 21
Db 16 VLLGGVL 22

RESULT 710
O59120_PYRHO PRELIMINARY; PRT; 617 AA.
AC O59120;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein PH1451.
GN OrderedLocuNames=PH1451;
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayasi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
```

```

RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Maeuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
EL DNA Res. 5:55-76(1998).
DR EMBL; BA000001; BAA30558.1; -; Genomic_DNA.
DR F1R; F71019; F71019.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR InterPro; IPR003918; NADH_oxred.
DR InterPro; IPR001750; Oxidored_q1.
DR Pfam; PF00361; Oxidored_q1; 1.
DR PRINTS; PR01437; NUOXDRDTASE4.
DR Complete proteome; Hypothetical protein.
DR SEQUENCE 617 AA; 67731 MW; 65B84AF5CCE04CAB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 617;
Best Local Similarity 100.0%; Pred. No. 7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 292 GGVLAAL 298

RESULT 711
Q9NXC4_HUMAN
ID Q9NXC4_HUMAN PRELIMINARY; PRT; 626 AA.
AC Q9NXC4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLN20327.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
CX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000334; BAA91091.1; -; mRNA.
DR Ensembl; ENSG00000147804; Homo sapiens.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0030001; F:metal ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003689; Zn_transp_zip.
DR Pfam; PF02535; Zip; 1.
DR SEQUENCE 626 AA; 66249 MW; 4F1E28B90ACA388C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 626;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 155 GGVLAAL 161

RESULT 712
Q4QAV0_LEIMA
ID Q4QAV0_LEIMA PRELIMINARY; PRT; 627 AA.
AC Q4QAV0;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)

```

```

DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=LmjP32.3220;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CT005269; CAJ08853.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 627 AA; 76399 MW; 7C3BEF99D09B8B1A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 EMERCSQ 67
Db 130 EMERCSQ 136

RESULT 713
Q5WB2_BACSK
ID Q5WB2_BACSK PRELIMINARY; PRT; 627 AA.
AC Q5WB2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN Ordered locus names=ABC3595;
OS Bacillus clausii (strain KSM-K16).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=66692;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KSM-K16;
RA Takai Y., Kageyama Y., Shimamura S., Suzuki H., Nishi S., Hatada Y.,
RA Kawai S., Ito S., Horikoshi K.;
RT "The complete genome sequence of the alkaliphilic Bacillus clausii
RT KSM-K16."
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006627; BAD66128.1; -; Genomic_DNA.
DR InterPro; IPR010144; Cas_CT1133.
DR TIGRFAMs; TIGR01863; cas_CT1133; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 627 AA; 72286 MW; 6A1C9CC595CB01AA CRC64;

Query Match 5.9%; Score 7; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 LGGKPAI 48
Db 174 LGGKPAI 180

RESULT 714
Q8VB5_MOUSE
ID Q8VB5_MOUSE PRELIMINARY; PRT; 627 AA.
AC Q8VB5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Evc protein.
GN Name=Evc;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Klausner R.D., Peingold E.A., Grouse L.H., Derge J.G.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.K., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT *Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences*;
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CZECH II;
 RC TISSUE=Mammary tumor metastasized to lung. Tumor arose spontaneously;
 RA Strausberg R.;
 RA Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC021464; AAH21464.1; -; mRNA.
 DR EMBL; BC021392; AAH21392.1; -; mRNA.
 DR MGI; MGI:1890596; Evc.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 SQ SEQUENCE 627 AA; 70253 MW; 01C197E258511EF0 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 627;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 15 VLLGGVL 21
 Db 27 VLLGGVL 33
 RESULT 715
 Q6LTG7 PHOPR
 ID Q6LTG7_PHOPR PRELIMINARY; PRT; 629 AA.
 AC Q6LTG7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein VP0151.
 GN Name=VP0151; OrderedLocNames=PBPR0998;
 OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Photobacterium.
 OX NCBI_TaxID=74109;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15746425; DOI=10.1126/science.1103341;
 RA Vezzi A., Campanaro S., D'Angelo M., Simonati F., Vitulo N.,
 RA Lauro P.M., Cestaro A., Malacrida G., Simonati B., Cannata N.,
 RA Rinaldi C., Bartlett D.H., Valle G.;
 RT *Life at depth: Photobacterium profundum genome sequence and
 RT expression analysis*;
 RL Science 307:1459-1461 (2005).
 RL EMBL; CR378666; CAG19409.1; -; Genomic_DNA.
 DR InterPro; IPR008503; DUF785.

DR InterPro; IPR002119; Histone_H2A.
 DR Pfam; PF05618; DUF785; 2.
 DR PROSITE; PS00046; HISTONE_H2A; UNKNOWN_1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 629 AA; 69888 MW; 0B95EA661E8451D2 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 629;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 71 YIEQAQV 77
 Db 497 YIEQAQV 503
 RESULT 716
 Q5B5C3 EMENI
 ID Q5B5C3 EMENI PRELIMINARY; PRT; 632 AA.
 AC Q5B5C3;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=AN4257.2;
 OS Aspergillus nidulans FGSC A4.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; Emericella.
 OX NCBI_TaxID=227321;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=FGSC A4;
 RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavsky L.,
 RA Bougaltier B., Butler J., Calvo S.E., Canarata J., Chang J.,
 RA Choquel Y., Collymore A., Cook A., Cooke P., Corum B., DeArelano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
 RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal A., Karatas A.,
 RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu J., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrum J., Meneus L.,
 RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT *Genome Sequence of Aspergillus nidulans*;
 RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACD01000069; EAA58925.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 632 AA; 70971 MW; 162346EB6F15ABF9 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 632;
 Best Local Similarity 100.0%; Pred. No. 7.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 17 LGGVLAA 23
 Db 597 LGGVLAA 603
 RESULT 717
 Q9RU98_DEIRA

```

ID AC Q9RU98 DEIRA PRELIMINARY; PRT; 645 AA.
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE NADH dehydrogenase I, L subunit.
GN OrderedLocustNames=DR1494;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Manathavan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.B., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AB011993; AAF11057.1; -; Genomic_DNA.
DR PIR; A75390; A75390.
DR TIGR; DR1494; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0008137; F:NADH dehydrogenase (ubiquinone) activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0042773; P:ATP synthesis coupled electron transport; IEA.
DR InterPro; IPR003945; NADHpl_oxrds.
DR InterPro; IPR003916; NADHpl_oxrds.
DR InterPro; IPR001750; Oxidored_q1.
DR InterPro; IPR001516; Oxidored_q1_N.
DR Pfam; PF00361; Oxidored_q1; 1.
DR Pfam; PF00662; Oxidored_q1_N; 1.
DR PRINTS; PR01434; NADHGNASE5.
DR PRINTS; PR01435; NPOXDRGNASE5.
DR TIGRFAMs; TIGR01974; NDH_I_L; 1.
DR Complete proteome; NAD; Oxidoreductase; Transmembrane.
SQ SEQUENCE 645 AA; 65945 MW; AB5A3493F4DDE3F6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 645;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAAL 25
Db 475 GVLAAAL 481

RESULT 718
Q9H6T8 HUMAN
ID Q9H6T8 HUMAN PRELIMINARY; PRT; 647 AA.
AC Q9H6T8
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ21884.
GN Name=SLC39A4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Oiyashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Director MGC Project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL; BC062625; AAH62625.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003689; Zn_transpt_Zip.
DR Pfam; PF02535; Zip; 1.

```

```

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025537; BAB15164.1; -; mRNA.
DR ENSEMBL; ENSG00000147804; Homo sapiens.
DR HGNC; HGNC:17129; SLC39A4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003689; Zn_transpt_Zip.
DR Pfam; PF02535; Zip; 1.
SQ SEQUENCE 647 AA; 68421 MW; BEB930F88BA1151F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 647;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 180 GGVLAAL 186

RESULT 719
Q6P5W5 HUMAN
ID Q6P5W5 HUMAN PRELIMINARY; PRT; 647 AA.
AC Q6P5W5;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Solute carrier family 39 (Zinc transporter), member 4.
GN Name=SLC39A4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Buetow K.H., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Lung;
RA Director MGC Project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL; BC062625; AAH62625.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR003689; Zn_transpt_Zip.
DR Pfam; PF02535; Zip; 1.

```

```

SQ SEQUENCE 647 AA; 68451 MW; 1EB287FE50CBA9A9 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 647;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 180 GGVLAAL 186

RESULT 720
Q5Z0R6 NOCPA
ID Q5Z0R6 NOCPA PRELIMINARY; PRT; 657 AA.
AC Q5Z0R6
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=nfall1300;
OS Nocardia farcinica.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Nocardia.
OX NCBI_TaxID=37329;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=IPM 10152;
RK PubMed=15466710; DOI=10.1073/pnas.0406410101;
RA Ishikawa J., Yamashita A., Mikami Y., Hoshino Y., Kurita H., Hotta K.,
RA Shiba T., Hattori M.;
RT "The complete genomic sequence of Nocardia farcinica IFM 10152." ;
DR EMBL; AP06618; BAD55975.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0017111; F:nucleoside-triphosphatase activity; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR SMART; SM00382; AAA; 1.
KW ATP-binding; Complete proteome; Hypothetical protein;
KW Nucleotide-binding; Transport.
SQ SEQUENCE 657 AA; 69569 MW; C2684F8132871AEC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 657;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 464 VLAALAA 470

RESULT 721
Q8ZBL8 YERPE
ID Q8ZBL8 YERPE PRELIMINARY; PRT; 665 AA.
AC Q8ZBL8; Q74XT8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Ferrichrome transport system permease protein Fthub.
GN Name=fthub; OrderedLocusNames=YPO295, YPO3330;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CO-92 / Biovar Orientalis;
RA MEDLINE=21470413; PubMed=11586360; DOI=10.1038/35097083;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C.M., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.-M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,

```

```

RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.M.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague." ;
RL Nature 413:523-527(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=91001;
RX PubMed=15368893;
RA Song Y., Tong Z., Wang J., Wang L., Guo Z., Han Y., Zhang J., Pei D.,
RA Zhou D., Qin H., Pang X., Han Y., Zhai J., Li M., Cui B., Qi Z.,
RA Jin L., Dai R., Chen F., Li S., Ye C., Du Z., Lin W., Wang J., Yu J.,
RA Yang H., Wang J., Huang P., Yang R.;
RT "Complete genome sequence of Yersinia pestis strain 91001, an isolate
RT avirulent to humans." ;
RL DNA Res. 11:179-197(2004).
DR EMBL; AJ414156; CAC92620.1; -; Genomic DNA.
DR EMBL; AE017128; AAS60570.1; -; Genomic_DNA.
DR PIR; AH0411; AH0411.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FecD.
DR Pfam; PF01032; FecCD; 2.
KW Complete proteome.
SQ SEQUENCE 665 AA; 70354 MW; DOBA062875CB0DC0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 665;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 399 VLAALAA 405

RESULT 722
Q66EF2 YERPS
ID Q66EF2 YERPS PRELIMINARY; PRT; 665 AA.
AC Q66EF2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE ABC type hydroxamate-dependent iron uptake, with duplicated permease
DE domains precursor.
GN Name=fthub; OrderedLocusNames=YPT80741;
OS Yersinia pseudotuberculosis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=633;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=IP32953 / Serotype I;
RX PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA Chain P.S.G., Carniel E., Larimer P.W., Lemerlin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the evolution of Yersinia pestis through whole-genome
RT comparison with Yersinia pseudotuberculosis." ;
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH19981.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FecD.
DR Pfam; PF01032; FecCD; 2.
KW Complete proteome; Signal.
FT SIGNAL 1 43 Potential.
SQ SEQUENCE 665 AA; 70354 MW; DOBA062875CB0DC0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 665;

```

```
Best Local Similarity 100.0%; Pred. No. 7.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 7; Conservative 0;

Qy 20 VLAALAA 26
Db 399 VLAALAA 405

RESULT 723
Q8ZUL8_PYRAE
ID Q8ZUL8_PYRAE PRELIMINARY; PRT; 667 AA.
AC Q8ZUL8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE Hypothetical protein PAF2717.
GN OrderedLocusNames=PAE2717;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989 (2002).
DR EMBL; AB009890; AAL64389.1; -; Genomic_DNA.
DR HSP; P80579; IQW.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR004879; DUF255.
DR Pfam; PF03190; DUF255; 1.
KW Complete proteome.
SQ SEQUENCE 667 AA; 74773 MW; 930F3257BC655113 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 667;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 185 VLAALAA 191

RESULT 724
Q584R6_9TRYP
ID Q584R6_9TRYP PRELIMINARY; PRT; 669 AA.
AC Q584R6;
DT 10-MAY-2005 (TRENBLrel. 30, Created)
DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
DE Hypothetical protein.
GN ORFNames=Tb927.6.2400;
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Ghedin E., Blandin G., Bartholomeu D., Caler E., Haas B., Hannick L.,
RA Shalton J., Hou L., Djikeng A., Feldblyum T., Hostetler J.,
RA Johnson J., Jones K., Koo H.L., Larkin C., Pai G., Peterson J.,
RA Khalak H.G., Salzberg S., Simpson A.J., Tallon L., Van Aken S.,
RA Wanless D., White O., Wortman J., Fraser C.M., El-Sayed N.M.A.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=GUTat10.1;
RA El-Sayed N.M., Khalak H., Adams M.D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=GUTat10.1;
RA Haas B., Blandin G., El-Sayed N.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC074259; AAX80872.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 669 AA; 74524 MW; 130F002EBE212BAD6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 566 VLAALAA 572

RESULT 725
HPPA_XANAC
ID HPPA_XANAC STANDARD; PRT; 675 AA.
AC Q8PH20;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Pyrophosphate-energized proton pump (EC 3.6.1.1) (Pyrophosphate-
DE energized inorganic pyrophosphatase) (H+-Pase) (Membrane-bound
DE proton-translocating pyrophosphatase).
GN Name=hppA; OrderedLocusNames=XAC3440;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chamberg F.,
RA Ciapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Gruber A.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463 (2002).
CC -!- FUNCTION: Generates a proton motive force; it probably catalyzes a
CC fully reversible reaction, thus being able to synthesize
CC pyrophosphate when the proton motive force is sufficient (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC -!- COPACTOR: Magnesium (By similarity).
CC -!- SUBUNIT: Homodimer (Potential).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the H(+) -translocating pyrophosphatase
CC (TC 3.A.10) family. Type 2 subfamily.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
```

```
CC removed.
CC -----
DR EMBL; AE011990; AM38283.1; -; Genomic_DNA.
DR HAMAP; MF_01130; -; 1.
DR InterPro; IPR004131; H_PPase.
DR PIRSF; PIRSF001265; H+-PPase; 1.
DR TIGRFAMs; TIGR01104; V_PPase; 1.
KW Complete proteome; Hydrogen ion transport; Hydrolase; Ion transport;
KW Magnesium; Transmembrane; Transport.
FT TRANSMEM 7 29 Potential.
FT TRANSMEM 56 78 Potential.
FT TRANSMEM 80 102 Potential.
FT TRANSMEM 127 149 Potential.
FT TRANSMEM 162 184 Potential.
FT TRANSMEM 231 250 Potential.
FT TRANSMEM 257 279 Potential.
FT TRANSMEM 285 307 Potential.
FT TRANSMEM 320 342 Potential.
FT TRANSMEM 368 390 Potential.
FT TRANSMEM 397 419 Potential.
FT TRANSMEM 458 480 Potential.
FT TRANSMEM 501 523 Potential.
FT TRANSMEM 564 586 Potential.
FT TRANSMEM 591 613 Potential.
FT TRANSMEM 652 674 Potential.
FT SITE 457 457 Determinant of potassium independence (By similarity).
SQ SEQUENCE 675 AA; 69285 MW; 05C2DAEDFA8106B CRC64;

Query Match 5.9%; Score 7; DB 1; Length 675;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 299 GVLAALA 305
|||||

RESULT 726
Q8A4X3 BACTN PRELIMINARY; PRT; 686 AA.
AC Q8A4X3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Cytochrome c biogenesis protein ccsA.
GN OrderedLocNames=BT2474;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RT Chiang H.C., Hooper L.V., Gordon J.I.;
RA "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016936; AA077581.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008335; P:cytochrome c oxidase complex assembly; IEA.
DR InterPro; IPR002541; CytC_asm.
DR PIRSF; PIRSF01578; Cytochrom_C_asm; 1.
KW Complete proteome.
SQ SEQUENCE 686 AA; 78318 MW; 5452BE1B9A056B7F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 686;
Best Local Similarity 100.0%; Pred. No. 7.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 299 GVLAALA 305
|||||
```

```
Db 51 GVLAALA 57
|||||

RESULT 727
Q8D1A4 YERPE PRELIMINARY; PRT; 689 AA.
AC Q8D1A4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Inner membrane permease of ABC transporter.
GN Name=fhbB, OrderedLocNames=y0798;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=KIM5 / Biovar. Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
RA Deng W., Butland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Petherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AE013682; AM84385.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FecD.
DR Pfam; PF01032; FecCD; 2.
SQ SEQUENCE 689 AA; 72974 MW; 2AED5C45EDF36609 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 689;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 423 VLAALAA 429
|||||

RESULT 728
Q9AF95_9RHO PRELIMINARY; PRT; 691 AA.
AC Q9AF95;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 1-butanol dehydrogenase BDH.
OS Pseudomonas butanovora.
OC Bacteria; Proteobacteria; Betaproteobacteria; Rhodocyclales;
OC Rhodocyclaceae; Thaueria.
OX NCBI_TaxID=86174;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC MEDLINE=21896383; PubMed=11889098;
RX DOI=10.1128/JB.184.7.1916-1924.2002;
RA Vangnai A.S., Arp D.J., Sayavedra-Soto L.A.;
RT "Two distinct alcohol dehydrogenases participate in butane metabolism
by Pseudomonas butanovora.";
RL J. Bacteriol. 184:1916-1924(2002).
DR EMBL; AF355798; AA027220.2; -; Genomic_DNA.
DR HSP; Q46444; 1KB0.
DR GO; GO:0020037; F:heme binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR01282; Cytochrome c R.
DR InterPro; IPR009056; Cyt_c_monohem.
DR InterPro; IPR002372; PQQ_repeat.
```



```

DR Pfam; PF01011; PQQ; 3.
DR SMART; SM00564; PQQ; 3.
DR PROSITE; PS51007; CYTC; 1.
SQ SEQUENCE 691 AA; 75070 MW; 4FC7PD20CDAL4E64 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 691;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 40 IELGGKP 46
Db 334 IELGGKP 340

RESULT 729
Q4LLY8_9BURK PRELIMINARY; PRT; 694 AA.
AC Q4LLY8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Ribonuclease II.
GN ORFNames=Bcen2424DRAFT.1832;
OS Burkholderia cenocepacia H12424.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; Burkholderia cepacia complex.
OX NCBI_TaxID=331272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
RT H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=H12424;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
RT H12424.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AALH01000047; EAM17085.1; -; Genomic DNA.
SQ SEQUENCE 694 AA; 75997 MW; D4C65DA41B76575E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 694;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGLL 90
Db 162 GKVLGLL 168

RESULT 730
Q6FTC9_CANGA PRELIMINARY; PRT; 704 AA.
AC Q6FTC9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to sp|P40071|Saccharomyces cerevisiae YER113c.
GN OrderedLocusNames=CAGL0G03487g;
OS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5478;
RN [1]

```

```

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 2001 / CBS 138;
RX PubMed=15229592; DOI=10.1038/nature02579;
RA LaFontaine I., Sherman D., Fischer G., Durrans P., Casaregola S.,
RA Lafont N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.-M., Boyne E., Bleykasten C.,
RA Boisarame A., Boyer J., Catolico L., Confanieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.-M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,
RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thiery A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Winkler P., Souciet J.-L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44 (2004).
DR EMBL; CR380953; CAG59442.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004240; EMP70.
DR PANTHER; PTHR10766; EMP70.
DR Pfam; PF02990; EMP70; 1.
KW Complete proteome.
SQ SEQUENCE 704 AA; 80688 MW; A60BD836D214519F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 704;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 22
Db 498 LGGVLA 504

RESULT 731
Q5WV32_LEGPL PRELIMINARY; PRT; 707 AA.
AC Q5WV32;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Guanosine-3',5'-bis(Diphosphate) 3'-pyrophosphohydrolase.
GN Name=Spot; OrderedLocusNames=lp11985;
OS Legionella pneumophila (strain Lens).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=297245;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15467720; DOI=10.1038/ng1447;
RA Cazalet C., Rusniok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
RA Tichit M., Jarraud S., Bouchier C., Vandenesch F., Kunst F.,
RA Etienne J., Glaser P., Buchrieser C.;
RT "Evidence in the Legionella pneumophila genome for exploitation of
RT host cell functions and high genome plasticity.";
RL Nat. Genet. 36:1165-1173 (2004).
DR EMBL; CR628337; CAH16225.1; -; Genomic DNA.
DR Legioli; lp11985; -.
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0015969; F:guanosine tetraphosphate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR006674; HD hydro.
DR InterPro; IPR003607; Met.phos hydro.
DR InterPro; IPR007685; RelA_SpOT.
DR InterPro; IPR004811; Spot_rela.
DR InterPro; IPR004095; TGS.
DR Pfam; PF01842; ACT; 1.

```

```

DR Pfam; PF01966; HD; 1.
DR Pfam; PF04607; RelA_Spot; 1.
DR Pfam; PF02824; TGS; 1.
DR SMART; SM00471; Hdc; 1.
DR TIGRFAMs; TIGR00691; spot_rela; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 707 AA; 79690 MW; E8A957871343279E CRC64;

Query Match 5.9%; Score 7; DB 2; Length 707;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
Db 640 GVLAALA 646

RESULT 732
Q5X3P4_LEGPA PRELIMINARY; PRT; 707 AA.
AC Q5X3P4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Guanosine-3', 5'-bis(Diphosphate) 3'-pyrophosphohydrolase.
GN Name=spot; OrderedLocuNames=Ipp1990;
OS Legionella pneumophila (strain Paris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=297246;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=13467720; DOI=10.1038/ngl1447;
RA Cazalot C., Rusniok C., Brueggemann H., Zidane N., Magnier A., Ma L.,
RA Ticht M., Jarraud S., Bucher C., Vandenesch F., Kunet F.,
RA Etienne J., Glaser P., Buchrieser C.;
RT "Evidence in the Legionella pneumophila genome for exploitation of
RT host cell functions and high genome plasticity.";
RL Nat. Genet. 36:1165-1173(2004).
RL EMBL; CR28336; CAH13142.1; -; Genomic_DNA.
DR Legiolist; Ipp1990; -
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0015969; P:guanosine tetraphosphate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR006674; HD hydro.
DR InterPro; IPR003607; Met_Phos hydro.
DR InterPro; IPR007685; RelA_Spot.
DR InterPro; IPR004811; Spot_rela.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF01966; HD; 1.
DR Pfam; PF04607; RelA_Spot; 1.
DR Pfam; PF02824; TGS; 1.
DR SMART; SM00471; Hdc; 1.
DR TIGRFAMs; TIGR00691; spot_rela; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 707 AA; 79748 MW; C6B44867204324AE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 707;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
Db 640 GVLAALA 646

RESULT 733
Q9K3K0_STRCO PRELIMINARY; PRT; 712 AA.
ID Q9K3K0_STRCO
AC Q9K3K0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative serine/threonine protein kinase.
GN OrderedLocuNames=SC04820; ORFNames=SC2A6.05c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinomycetaceae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=A3(2) / M145;
MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL393121; CAB97423.1; -; Genomic_DNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; Kinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
KW ATP-binding; Complete proteome; Kinase; Nucleotide-binding;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 712 AA; 74552 MW; F116E08EBDC8694 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 712;
Best Local Similarity 100.0%; Pred. No. 7.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 507 VLAALAA 513

RESULT 734
Q4HHU9_CAMCO PRELIMINARY; PRT; 713 AA.
ID Q4HHU9_CAMCO
AC Q4HHU9;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE [NifE] hydrogenase maturation protein HypF.
GN Name=HypF; ORFNames=CC00704;
OS Campylobacter coli RM2228.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=306254;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=RM2228;
RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
RA Nelson K.E.;
RT "Major structural and novel potential virulence mechanisms from the
RT genomes of multiple Campylobacter species.";
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```

```

CC preliminary data.
DR EMBL; AAFJ01000001; EAL57436.1; -: Genomic DNA.
SQ SEQUENCE 713 AA; 18178 MW; 9386CGDFOCFBI04A CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7.9e+02; Length 713;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 258 LLGGVLA 264

RESULT 735
Q72FC0 DESVH
ID Q72FC0_DESVH PRELIMINARY; PRT; 714 AA.
AC Q72FC0;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Glycosyl transferase, group 2 family protein.
GN OrderedLocusNames=DVU0294;
OS Desulfobrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB 8303)
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobirionales;
OC Desulfobirionaceae; Desulfobrio.
OX NCBI_TaxID=882;
[1]
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Henne C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dmitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum I.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfobrio vulgaris Hildenborough."
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AE017310; AAS94777.1; -: Genomic DNA.
DR TIGR; DVU0294; -.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR001173; Glyco trans. 2.
DR Pfam; PF00535; Glycos transf. 2; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 714 AA; 76349 MW; 1129B6DD9CB61FD3 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7.9e+02; Length 714;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAAALAY 27
Db 293 LAAALAY 299

RESULT 736
Q5ZT29 LEGPH
ID Q5ZT29_LEGPH PRELIMINARY; PRT; 715 AA.
AC Q5ZT29;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Guanosine-3, 5-bis(diphosphate)-3-pyrophosphohydrolase
DE (EC 3.1.7.2).
GN Name=spot; OrderedLocusNames=lpq2009;
GN Legionella pneumophila subsp. pneumophila (strain Philadelphia 1 /
OS ATCC 33152).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Legionellaceae; Legionella.
OX NCBI_TaxID=272624;
[1]
RN NUCLEOTIDE SEQUENCE.
RX PubMed=15448271; DOI=10.1126/science.1099776;
RA Chien M., Morozova I., Shi S., Sheng H., Chen J., Gomez S.M.,
RA Asamari G., Hill K., Nuara J., Feder M., Rineer J., Greenberg J.J.,
RA Steshenko V., Park S.H., Zhao B., Teplickaya E., Edwards J.R.,
RA Pampou S., Georgiou A., Chou I.-C., Iannuccilli W., Ulz M.E.,
RA Kim D.H., Geringer-Sameth A., Goldsberry C., Morozov P., Fischer S.G.,
RA Segal G., Qu X., Rzhetsky A., Zhang P., Cavanis E., De Jong P.J.,
RA Ju J., Kalachikov S., Shuman H.A., Russo J.J.;
RT "The genomic sequence of the accidental pathogen Legionella
RT pneumophila";
RL Science 305:1966-1968(2004).
DR EMBL; AB017354; AAU28078.1; -: Genomic DNA.
DR GO; GO:0016597; F:amino acid binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0015969; P:guanosine tetraphosphate metabolism; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR002912; ACT.
DR InterPro; IPR006674; HD hydro.
DR InterPro; IPR003607; Met_phos hydro.
DR InterPro; IPR007685; RelA Spot.
DR InterPro; IPR004811; Spot_rela.
DR InterPro; IPR004095; TGS.
DR Pfam; PF01842; ACT; 1.
DR Pfam; PF01966; HD; 1.
DR Pfam; PF04607; RelA Spot; 1.
DR Pfam; PF02824; TGS; 1.
DR SMART; SM00471; HDC; 1.
DR TIGRPFAMs; TIGR00691; spot_rela; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 715 AA; 80670 MW; 9E3E98F07EF5AD81 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7.9e+02; Length 715;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAALA 25
Db 648 GVLAALA 654

RESULT 737
Q4NT17_9DELT
ID Q4NT17_9DELT PRELIMINARY; PRT; 728 AA.
AC Q4NT17;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=AdehDRAFT1807;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cytophactereae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
[1]
RN NUCLEOTIDE SEQUENCE.
RX STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

```



```

Qy 16 LLGGVLA 22
Db 274 LLGGVLA 280

RESULT 741
Q9PHF5 CAMJE PRELIMINARY; PRT; 729 AA.
AC Q9PHF5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Transcriptional regulatory protein hypF.
GN Name=hypF; OrderedLocusNames=CJ0622;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NCCT 11169;
RX MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;
RA Parkhill J., Wren B.W., Mungall K.L., Ketley J.M., Churcher C.M.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagsels K., Kariyeh A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrell B.G.;
RA "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
DR EMBL; AL139075; CAB75258.1; -; Genomic_DNA.
DR FIR; E81410; E81410.
DR HSSP; P30131; 1GXU.
DR GO; GO:0003998; P:acetylphosphatase activity; IEA.
DR GO; GO:0008450; P:O-sialoglycoprotein endopeptidase activity; IEA.
DR GO; GO:0030528; F:transcription regulator activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001792; Acylphosphatase.
DR InterPro; IPR004421; HyPF.
DR InterPro; IPR000905; Peptidase M22.
DR InterPro; IPR006070; Sua5/YciO/YrDC_N.
DR InterPro; IPR011125; Znf HYPF.
DR Pfam; PF00708; Acylphosphatase; 1.
DR Pfam; PF00814; Peptidase M22; 1.
DR Pfam; PF01300; Sua5_YciO_YrDC; 1.
DR Pfam; PF07503; zf-HYPF; 2.
DR PIRSF; PIRSF006256; CMPcnv_hdrg_mat; 1.
DR ProDom; PD001884; Acylphosphatase; 1.
DR TIGRFAMs; TIGR00143; hyPF; 1.
DR PROSITE; PS00150; ACYLPHOSPHATASE_1; 1.
KW Complete proteome.
SQ SEQUENCE 729 AA; 83507 MW; 9D6546AB154059DC CRC64;

Query Match 5.9%; Score 7; DB 2; Length 729;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 274 LLGGVLA 280

RESULT 742
Q92XH9 RHIME PRELIMINARY; PRT; 733 AA.
AC Q92XH9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Probable oxidoreductase.
GN OrderedLocusNames=RA1273; ORFNames=SMa2353;

Rhizobium meliloti (Sinorhizobium meliloti).
Plasmid pSymA.
Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
NCBI_TaxID=382;
[1]
NUCLEOTIDE SEQUENCE.
STRAIN=1021;
MEDLINE=21396509; PubMed=11481432; DOI=10.1073/pnas.161294798;
Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.
Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
Kaman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
"Nucleotide sequence and predicted functions of the entire
Sinorhizobium meliloti pSymA megaplasmid.";
Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
EMBL; AB07312; AAK6593.1; -; Genomic_DNA.
FIR; A95421; A95421.
GO; GO:0016491; F:oxidoreductase activity; IEA.
GO; GO:0006118; P:electron transport; IEA.
InterPro; IPR008274; Aldxan_dh_bind.
InterPro; IPR000674; Aldxan_dh_hamm.
InterPro; IPR012219; CO_dh_molybdo.
Pfam; PF01315; Ald_Xan_dh_C; 1.
Pfam; PF02738; Ald_Xan_dh_C2; 1.
PIRSF; PIRSF00129; CO_dh_molybdo; 1.
KW Complete proteome; Plasmid.
SQ SEQUENCE 733 AA; 78368 MW; FE37F41C012520B6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 733;
Best Local Similarity 100.0%; Pred. No. 8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
Db 253 VLALAA 259

RESULT 743
Q4UWV7_XANCP PRELIMINARY; PRT; 735 AA.
ID Q4UWV7_XANCP PRELIMINARY;
AC Q4UWV7;
DT 13-SEP-2005 (TReMBLrel. 31, Created)
DT 13-SEP-2005 (TReMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE Oxidoreductase.
GN ORFNames=XC_1387;
OS Xanthomonas campestris pv. campestris str. 8004.
Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000050; AAY48456.1; -; Genomic_DNA.
SQ SEQUENCE 735 AA; 77812 MW; 1B887187205108D6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 735;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLALAA 26
Db 254 VLALAA 260

```

```

RESULT 744
Q8P786_XANCP PRELIMINARY; PRT; 735 AA.
AC Q8P786;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN Name=yagr;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
Camargo L.E.A., Camarotte G., Cannavan F., Cardoso J., Chambergo F.,
Clapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
Ferreiro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
Satubal J.C., Kitajima J.P.;
"Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
REMBL; AB012385; FMM41999.1; -; Genomic DNA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008274; Aldxan_dh bind.
DR InterPro; IPR012219; CO_dh_molybdo.
DR Pfam; PF01315; Ald_xan_dh_C2; 1.
DR Pfam; PF02738; Ald_xan_dh_C2; 1.
DR PIRSF; PIRSF000129; CO_dh_molybdo; 1.
KW Complete proteome.
SQ SEQUENCE 735 AA; 77811 MW; 1B887187205108D6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 735;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 254 VLAALAA 260
|||||

RESULT 745
Q4UQ53_XANCP PRELIMINARY; PRT; 737 AA.
AC Q4UQ53;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Oxidoreductase.
GN ORFNames=XC 3780;
OS Xanthomonas campestris pv. campestris str. 8004.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=314565;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RX Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-P.,

```

```

RA Sun Q.-H., Ying G., Tang D.-J., Wu W., Wang L.-P., Jiang B.-L.,
Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000050; AAY50820.1; -; Genomic DNA.
SQ SEQUENCE 737 AA; 78034 MW; 43A4AB1BFD980510 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 737;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 257 VLAALAA 263
|||||

RESULT 746
Q8P4J8_XANCP PRELIMINARY; PRT; 737 AA.
AC Q8P4J8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Oxidoreductase.
GN Name=yagr;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=22022145; PubMed=12024217; DOI=10.1038/417459a;
da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furian L.R.,
Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
Camargo L.E.A., Camarotte G., Cannavan F., Cardoso J., Chambergo F.,
Clapina L.P., Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
El-Dorri H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
Ferreiro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
Satubal J.C., Kitajima J.P.;
"Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
REMBL; AB012491; FMM42967.1; -; Genomic DNA.
DR GO; GO:0016491; P:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008274; Aldxan_dh bind.
DR InterPro; IPR000674; Aldxan_dh hamm.
DR Pfam; PF01315; Ald_xan_dh_C2; 1.
DR Pfam; PF02738; Ald_xan_dh_C2; 1.
DR PIRSF; PIRSF000129; CO_dh_molybdo; 1.
KW Complete proteome.
SQ SEQUENCE 737 AA; 78034 MW; 43A4AB1BFD980510 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 737;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 257 VLAALAA 263
|||||

```

```

RESULT 747
ID 083118 TREPA PRELIMINARY; PRT; 740 AA.
AC 083118;
DT 01-NOV-1998 (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein TP0079.
GN OrderedLocusNames=TP0079;
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876; DOI=10.1126/science.281.5375.375;
RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RA Dodson R.J., Gwinn M.L., Hickey E.K., Clayton R.A., Ketchum K.A.,
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S.L., Peterson J.D.,
RA Khalak H.G., Richardson D.L., Howell J.K., Chidambaram M.,
RA Utterback T.R., McDonald L.A., Artach P., Bowman C., Cotton M.D.,
RA Fujii C., Garland S.A., Hatch B., Horst K., Roberts K.M., Sandusky M.,
RA Weidman J.F., Smith H.O., Venter J.C.;
RT "Complete genome sequence of Treponema pallidum, the syphilis
RT spirochete.";
RL Science 281:375-388 (1998).
DR EMBL; AB001192; AAC65074.1; -; Genomic_DNA.
DR PIR; F71369; F71369.
DR TIGR; TP0079; -.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR008274; Aldxan_dh_bind.
DR InterPro; IPR000674; Aldxan_dh_hamm.
DR Pfam; PF01315; Ald_Xan_dh_C1; 1.
DR Pfam; PF02738; Ald_Xan_dh_C2; 1.
DR Complete proteome.
KW SEQUENCE 740 AA; 81778 MW; 7D8AE3DEB4D22ED6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 740;
Best Local Similarity 100.0%; Pred. No. 8.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAYC 28
Db 242 AALAAAYC 248
|||||

RESULT 748
ID Q51AK4 ENTHI PRELIMINARY; PRT; 755 AA.
AC Q51AK4;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Cleavage and polyadenylation specificity factor 73 kDa subunit,
DE putative.
DE ORFNames=33.t00006;
GN Entamoeba histolytica HM-1:IMSS.
OS Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=294381;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HM-1:IMSS;
RX PubMed=15729342; DOI=10.1038/nature03291;
RA Loftus B., Anderson I., Davies R., Alsmark U.C., Samuelson J.,
RA Amedeo P., Roncaglia P., Berriman M., Hirt R.P., Mann B.J., Nozaki T.,
RA Suh B., Pop M., Duchene M., Ackers J., Tannich E., Leippe M.,
RA Hofer M., Bruchhaus I., Willhoeft U., Bhattacharya A.,
RA Chillingworth T., Churcher C., Hance Z., Harris B., Harris D.,
RA Jagels K., Moule S., Mungall K., Ormond D., Squares R., Whitehead S.,
RA Quail M.A., Rabinowitz E., Norbertczak H., Price C., Wang Z.,
RA Guillen N., Gilchrist C., Stroup S.E., Bhattacharya S., Lohia A.,
RA Foster P.G., Sicheritz-Ponten T., Weber C., Singh U., Mukherjee C.,

```

```

RA El-Sayed N.M., Petri W.A., Clark C.G., Embley T.M., Barrell B.,
RA Fraser C.M., Hall N.;
RT "The genome of the protist parasite Entamoeba histolytica.";
RL Nature 433:865-868(2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAFB01000138; EAL49856.1; -; Genomic DNA.
DR EMBL; AAFB01000138; EAL49856.1; -; Genomic DNA.
SQ SEQUENCE 755 AA; 86770 MW; 54EA544AE6210DB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 755;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKVL 87
Db 98 QFKGKVL 104
|||||

RESULT 749
ID Q7NDC5 GLOVI PRELIMINARY; PRT; 758 AA.
AC Q7NDC5;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE GLr4310 protein.
GN OrderedLocusNames=glr4310;
OS Gloeobacter violaceus.
OC Bacteria; Cyanobacteria; Gloeobacteriales; Gloeobacter.
OX NCBI_TaxID=33072;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=PCC 7421;
RX MEDLINE=22977040; PubMed=14621292;
RA Nakamura Y., Kaneko T., Sato S., Mimuro M., Miyashita H., Tsuchiya T.,
RA Sasamoto S., Watanabe A., Kawashima K., Kishida Y., Kiyokawa C.,
RA Kohara M., Matsumoto M., Matsuno A., Nakazaki N., Shimpō S.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of Gloeobacter violaceus PCC 7421, a
RT cyanobacterium that lacks thylakoids.";
RL DNA Res. 10:137-145(2003).
DR EMBL; BA000045; BAC92251.1; -; Genomic_DNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0009103; P:lipopolysaccharide biosynthesis; IEA.
DR InterPro; IPR003856; LPS_Wzz_MPA.
DR Pfam; PF02706; Wzz; 1.
DR Complete proteome.
KW SEQUENCE 758 AA; 82796 MW; B77AB1BB0C4FE469 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 758;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAAAY 27
Db 457 LAALAAAY 463
|||||

RESULT 750
ID Q8C4W5 MOUSE PRELIMINARY; PRT; 766 AA.
AC Q8C4W5;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
DE enriched library, clone:A73008A17 product:hypothetical ARM repeat
DE structure containing protein, full insert sequence. (Fragment).
GN Name=1810009A16rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

```

OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staehli P., Suzuki R., Tonita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC THE PANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=20493374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu N., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto K., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=C57BL/6J; TISSUE=Cerebellum;
RC Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hangaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
RA Hori P., Imotani K., Ishii Y., Itoh M., Kogawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,

RA Nishii K., Nomura K., Numazaki R., Ohno M., Oheato N., Okazaki Y.,
RA Saito R., Saichoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK080556; BAC37944.1; -; mRNA.
DR MGI; MGI:1916366; 1810009A16Rik.
KW Hypothetical protein.
FT NON TER 1 766 766
FT NON TER 1 766 766
SQ SEQUENCE 766 AA; 81870 MW; E2A082C0636728AF CRC64;

Query Match 5.9%; Score 7; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 8.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qv 20 VLAALAA 26
| | | | |
Db 731 VLAALAA 737

RESULT 751
Q8PQ37_XANAC PRELIMINARY; PRT; 769 AA.
ID Q8PQ37_XANAC PRELIMINARY;
AC Q8PQ37;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Two-component system sensor protein.
GN OrderedLocusNames=XAC0494;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX STRAIN=306 / ATCC 13902 / XV 101;
RC MEDLINE=22024217; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Canarotte G., Cannavan P., Cardozo J., Chambergo P.,
RA Clapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White P.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AB011676; AAM35383.1; -; Genomic_DNA.
DR HSP; P39928; 10XK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003677; P:DNA binding; IEA.
DR GO; GO:0016301; F:Kinase activity; IEA.
DR GO; GO:0000156; F:two-component response regulator activity; IEA.
DR GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0000160; P:two-component signal transduction system (p. . .); IEA.
DR InterPro; IPR003594; ATPbind_ATPase.
DR InterPro; IPR005467; His_kinase.
DR InterPro; IPR003661; His_kin_N.
DR InterPro; IPR004358; His_kin_like_C.
DR InterPro; IPR008207; Hpt.
DR InterPro; IPR001789; Response_reg.


```

DR Pfam: PF02518; HATPase_c; 1.
DR Pfam: PF00512; HsKA; 1.
DR Pfam: PF01627; Hpt; 1.
DR Pfam: PF00072; Response reg; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR ProDom: PD000039; Response reg; 1.
DR SMART: SM00387; HATPase_c; 1.
DR SMART: SM00388; HsKA; 1.
DR SMART: SM00073; HPT; 1.
DR SMART: SM00448; REC; 1.
DR PROSITE: PS0109; HIS_KIN; 1.
DR PROSITE: PS0894; HPT; 1.
DR PROSITE: PS0110; RESPONSE_REGULATORY; 1.
KW Complete proteome.
SQ SEQUENCE 769 AA; 83699 MW; B7995CA40A787593 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 16 VLAALAA 22

RESULT 752
Q8AAK5 BACTN
ID Q8AAK5 BACTN PRELIMINARY; PRT; 774 AA.
AC
AD
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Beta-hexosaminidase.
GN OrderedLocusNames=BT0459;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
ON NCBI_TaxID=818;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL: AE016927; AA075566.1; -: Genomic_DNA.
DR HSPF: O85361; IHP4.
DR GO: GO:0004563; F:beta-N-acetylhexosaminidase activity; IEA.
DR GO: GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO: GO:0005975; P:carbohydrate metabolism; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR InterPro: IPR000421; FA58_C.
DR InterPro: IPR001540; Glyco_hydro_20.
DR Pfam: PF00754; F5_F8_type_C; 1.
DR Pfam: PF00728; Glyco_hydro_20; 1.
DR PRINTS: PR00738; GLHYDRLAS20.
DR PROSITE: PS00022; FA58C_3; 1.
KW Complete proteome.
SQ SEQUENCE 774 AA; 86488 MW; DC00782C4541022C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 27
Db 271 LAALAA 277

RESULT 753
Q7WAE7 BORPA
ID Q7WAE7 BORPA PRELIMINARY; PRT; 774 AA.
AC
AD
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cation-transporting ATPase.
GN OrderedLocusNames=BP1430;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
ON NCBI_TaxID=519;
OX [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Davis P., Doggett J.,
RA Chillingworth T., Collins M., Cronin A., Hain P., Oggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL: BX640427; CAE36732.1; -: Genomic_DNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO: GO:0003824; F:catalytic activity; IEA.
DR GO: GO:0005261; F:channel activity; IEA.
DR GO: GO:0016818; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO: GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO: GO:0046872; F:metal ion binding; IEA.
DR GO: GO:0046873; F:metal ion transporter activity; IEA.
DR GO: GO:0008152; F:metabolism; IEA.
DR GO: GO:0030001; P:metal ion transport; IEA.
DR InterPro: IPR006416; ATPase-IB_hvy.
DR InterPro: IPR001757; ATPase-IB_E2.
DR InterPro: IPR001366; Cad_ATPase.
DR InterPro: IPR005834; Denal_like_hydro.
DR InterPro: IPR008250; E1-E2_ATPase_reg.
DR InterPro: IPR006404; Heavy_met_ATPase.
DR InterPro: IPR006121; Heavy_met_transpt.
DR InterPro: IPR000150; Hypothet_cof.
DR Pfam: PF00122; E1-E2_ATPase; 1.
DR Pfam: PF00403; HMA; 1.
DR Pfam: PF00702; Hydrolase; 1.
DR PRINTS: PR00119; CATATPASE.
DR PRINTS: PR00941; CDATPASE.
DR TIGRfams: TIGR01512; ATPase-IB2_Cd; 1.
DR TIGRfams: TIGR01525; ATPase-IB_hvy; 1.
DR TIGRfams: TIGR01494; ATPase_P-type; 2.
DR PROSITE: PS00154; ATPase_E1_E2; UNKNOWN_1.
DR PROSITE: PS01229; COF_2; UNKNOWN_1.
DR PROSITE: PS0846; HMA_2; 1.
KW ATP-binding; Complete proteome; Hydrolase; Nucleotide-binding;
KW Transmembrane.
SQ SEQUENCE 774 AA; 80722 MW; 26A3C74548D4BB26 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLAAL 24
Db 158 GGVLAAL 164

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 754
Q7WJ0 BORR
ID Q7WJ0 BORR PRELIMINARY; PRT; 778 AA.
AC Q7WJ0;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cation-transporting ATPase.
GN OrderedLocuNames=BB2504;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=RB50 / ATCC BAA-588;
RC MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; BX640444; CAB32998.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0015662; F:ATPase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005261; F:cation channel activity; IEA.
DR GO; GO:0016818; F:hydrolase activity, acting on acid anhydrid. ; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. ; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR006416; ATPase-IB hvy.
DR InterPro; IPR001366; Cad_ATPase.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2 ATPase reg.
DR InterPro; IPR006404; Heavy met ATPase.
DR InterPro; IPR006121; HeavyMe transpt.
DR InterPro; IPR000150; Hypothet_cof.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00403; HMA; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00941; CDATPASE.
DR TIGRFAMs; TIGR01512; ATPase-IB2 Cd; 1.
DR TIGRFAMs; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 2.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN_1.
DR PROSITE; PS01229; COF_2; UNKNOWN_1.
DR PROSITE; PS00846; HMA_2; 1.
KW ATP-binding; Complete proteome; Hydrolase; Nucleotide-binding;
KW Transmembrane.
SQ SEQUENCE 778 AA; 81192 MW; 76C20F43FD4D23A8 CRC64;

```

Query Match 5.9%; Score 7; DB 2; Length 778;
 Best Local Similarity 100.0%; Pred. No. 8.4e+02;

```

RESULT 755
Q7VVVB BORPE
ID Q7VVVB BORPE PRELIMINARY; PRT; 778 AA.
AC Q7VVVB;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cation-transporting ATPase.
GN OrderedLocuNames=BP2722;
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
NUCLEOTIDE SEQUENCE.
RC STRAIN=Tohana I / ATCC BAA-589 / NCTC 13251;
RC MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtmann M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; BX640419; CAB32999.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0015662; F:ATPase activity; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0005261; F:cation channel activity; IEA.
DR GO; GO:0016818; F:hydrolase activity, acting on acid anhydrid. ; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. ; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0046873; F:metal ion transporter activity; IEA.
DR GO; GO:0008152; F:metabolism; IEA.
DR GO; GO:0030001; P:metal ion transport; IEA.
DR InterPro; IPR006416; ATPase-IB hvy.
DR InterPro; IPR001366; Cad_ATPase.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2 ATPase reg.
DR InterPro; IPR006404; Heavy met ATPase.
DR InterPro; IPR006121; HeavyMe transpt.
DR InterPro; IPR000150; Hypothet_cof.
DR Pfam; PF00122; E1-E2 ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00941; CDATPASE.
DR TIGRFAMs; TIGR01512; ATPase-IB2 Cd; 1.
DR TIGRFAMs; TIGR01525; ATPase-IB_hvy; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 2.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN_1.
DR PROSITE; PS01229; COF_2; UNKNOWN_1.
KW ATP-binding; Complete proteome; Hydrolase; Nucleotide-binding;
KW Transmembrane.
SQ SEQUENCE 778 AA; 81176 MW; 5E6CE7F58CB8EA2C CRC64;

```

```

Query Match      5.9%; Score 7; DB 2; Length 778;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      18 GGVLAAL 24
Db      163 GGVLAAL 169

RESULT 756
QBPNP6_CORGL PRELIMINARY; PRT; 785 AA.
AC QBPNP6;
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Superfamily I DNA and RNA helicases (EC 3.6.1.-).
GN OrderedLocusNames=Cg11765;
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BA000036; BAB99156.1; -; Genomic_DNA.
DR HSSP; P56255; 1PJR.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004003; F:ATP-dependent DNA helicase activity; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003724; F:RNA helicase activity; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR InterPro; IPR000212; UvrD-helicase.
DR InterPro; IPR000806; Viral helicase.
DR Pfam; PF00580; UvrD-helicase; 2.
DR Pfam; PF01443; Viral helicase; 1.
KW Complete proteome; Helicase; Hydrolase.
SQ SEQUENCE 785 AA; 87594 MW; AFOB1AE541320C94 CRC64;

Query Match      5.9%; Score 7; DB 2; Length 785;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      5 ADLEVT 11
Db      714 ADLEVT 720

RESULT 757
Q742J2_MYCPA PRELIMINARY; PRT; 790 AA.
AC Q742J2;
DT 05-JUL-2004 (TREMELrel. 27, Created)
DT 05-JUL-2004 (TREMELrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE CtpB.
GN Name=ctpB; OrderedLocusNames=MAP0843;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium avium complex (MAC).
OX NCBI_TaxID=1770;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=k10;

Query Match      5.9%; Score 7; DB 2; Length 790;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      20 VLALAA 26
Db      324 VLALAA 330

RESULT 758
CTPE_MYCBO STANDARD; PRT; 797 AA.
AC P0A505; O08365;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable cation-transporting ATPase B (EC 3.6.3.-).
GN Name=ctpB; OrderedLocusNames=Mb0932;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1765;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garnier T., Biglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RL "The complete genome sequence of Mycobacterium bovis.";
Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type) family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; BX248337; CAD93793.1; -; Genomic_DNA.

```

```

DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR PANTHER; PTHR11939; ATPase_E1-E2; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 1.
DR PROSITE; PS00154; ATPase_E1-E2; 1.
KW ATP-binding; Complete proteome; Hydrolase; Magnesium; Metal-binding;
KW Nucleotide-binding; Phosphorylation; Transmembrane.
FT TRANSMEM 55 75 Potential.
FT TRANSMEM 215 235 Potential.
FT TRANSMEM 254 274 Potential.
FT TRANSMEM 549 569 Potential.
FT TRANSMEM 601 621 Potential.
FT TRANSMEM 633 653 Potential.
FT TRANSMEM 667 687 Potential.
FT TRANSMEM 703 723 Potential.
FT TRANSMEM 729 749 Potential.
FT TRANSMEM 764 784 Potential.
FT ACT_SITE 301 301 4-aspartylphosphate intermediate (By
FT METAL 536 536 Magnesium (By similarity).
FT METAL 540 540 Magnesium (By similarity).
SQ SEQUENCE 797 AA; 84973 MW; 4C5034FC6052FC7B CRC64;

Query Match 5.9%; Score 7; DB 1; Length 797;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 331 VLAALAA 337

RESULT 759
CTPE MYCTU STANDARD; PRT; 797 AA.
AC POA504; O08365;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Probable cation-transporting ATPase B (EC 3.6.3.-).
GN Namesctpb; OrderedLocuNames=Rv0908, MT0931; ORFNames=MTCY21C12.02;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC Mycobacterium tuberculosis complex.
OX NCBI_TaxID=1773;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jorgensen K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=22206494; PubMed=12218036;
RX DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

```

```

RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Onayama M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
laboratory strains.";
RL J. Bacteriol. 184:5479-5490 (2002).
CC -|- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- SIMILARITY: Belongs to the cation transport ATPase (P-type)
family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC -----
DR EMBL; BX842574; CAB08506.1; -; Genomic_DNA.
DR EMBL; AE000516; AAK45178.1; -; Genomic_DNA.
DR PIR; D70581; D70581.
DR TIGR; MT0931; -.
DR Tuberculist; Rv0908; -.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR PANTHER; PTHR11939; ATPase_E1-E2; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 1.
DR PROSITE; PS00154; ATPase_E1-E2; 1.
KW ATP-binding; Complete proteome; Hydrolase; Magnesium; Metal-binding;
KW Nucleotide-binding; Phosphorylation; Transmembrane.
FT TRANSMEM 55 75 Potential.
FT TRANSMEM 215 235 Potential.
FT TRANSMEM 254 274 Potential.
FT TRANSMEM 549 569 Potential.
FT TRANSMEM 601 621 Potential.
FT TRANSMEM 633 653 Potential.
FT TRANSMEM 667 687 Potential.
FT TRANSMEM 703 723 Potential.
FT TRANSMEM 729 749 Potential.
FT TRANSMEM 764 784 Potential.
FT ACT_SITE 301 301 4-aspartylphosphate intermediate (By
FT METAL 536 536 Magnesium (By similarity).
FT METAL 540 540 Magnesium (By similarity).
FT CONFLICT 736 736 A -> P (in Ref. 2).
SQ SEQUENCE 797 AA; 84973 MW; 4C5034FC6052FC7B CRC64;

Query Match 5.9%; Score 7; DB 1; Length 797;
Best Local Similarity 100.0%; Pred. No. 8.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 331 VLAALAA 337

RESULT 760
SL9A1_HUMAN STANDARD; PRT; 815 AA.
AC P19634;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1) (Solute
carrier family 9 member 1) (Na(+)/H(+) antiporter, amiloride-
sensitive) (APNH).
DE -----

```

GN Name=SLC9A1; Synonyms=APNH1, NHE1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC MEDLINE=89106219; PubMed=2536298; DOI=10.1016/0092-8674(89)90901-X;
RX Sartet C., Franchi A., Pouyssegur J.;
RA "Molecular cloning, primary structure, and expression of the human
RT growth factor-activatable Na⁺/H⁺ antiporter.";
RL Cell 56:271-280(1989).
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Kidney;
RC MEDLINE=90140739; PubMed=2154036;
RX Sartet C., Counillon L., Franchi A., Pouyssegur J.;
RA "Growth factors induce phosphorylation of the Na⁺/H⁺ antiporter,
RT glycoprotein of 110 kD.";
RL Science 247:723-726(1990).
RN [3]
RN SEQUENCE REVISION.
RX MEDLINE=91293066; PubMed=1712287;
RX Tse C.-W., Ma A.I., Yang V.W., Watson A.J.M., Levine S.,
RA Montrose M.H., Potter J., Sartet C., Pouyssegur J., Donowitz M.;
RA "Molecular cloning and expression of a cDNA encoding the rabbit ileal
RT villus cell basolateral membrane Na⁺/H⁺ exchanger.";
RL EMO J. 10:1957-1967(1991).
RN [4]
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Heart;
RC MEDLINE=94111706; PubMed=8283968;
RX Fliegel L., Dyck J.R., Wang H., Fong C., Haworth R.S.;
RA "Cloning and analysis of the human myocardial Na⁺/H⁺ exchanger.";
RL Mol. Cell. Biochem. 125:137-143(1993).
RN [5]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=20375279; PubMed=10913675; DOI=10.1016/S0165-4608(99)00246-0;
RX Garden O.A., Musk P., Worthington-White D.A., Dewey M.J., Rich I.N.;
RA "Silent polymorphisms within the coding region of human
RT sodium/hydrogen exchanger isoform-1 cDNA in peripheral blood
RT mononuclear cells of leukemia patients: a comparison with healthy
RL controls.";
RL Cancer Genet. Cytogenet. 120:37-43(2000).
RN [6]
RN INTERACTION WITH CHP.
RX MEDLINE=21248790; PubMed=11350981; DOI=10.1074/jbc.M100296200;
RA Pang T., Su X., Wakabayashi S., Shigekawa M.;
RT "Calcineurin homologous protein as an essential cofactor for Na⁺/H⁺
RT exchangers.";
RL J. Biol. Chem. 276:17367-17372(2001).
RN [7]
RN INTERACTION WITH CHP2.
RX MEDLINE=22313590; PubMed=12226101; DOI=10.1074/jbc.M208313200;
RA Pang T., Wakabayashi S., Shigekawa M.;
RT "Expression of calcineurin B homologous protein 2 protects serum
RT deprivation-induced cell death by serum-independent activation of
RT Na⁺/H⁺ exchanger.";
RL J. Biol. Chem. 277:43771-43777(2002).
CC CC -1- FUNCTION: Involved in pH regulation to eliminate acids generated
CC by active metabolism or to counter adverse environmental
CC conditions. Major proton extruding system driven by the inward
CC sodium ion chemical gradient. Plays an important role in signal
CC transduction.
CC CC -1- BIOPHYSICOCHEMICAL PROPERTIES:
CC pH dependence:
CC Fully active at acidic pHs, the antiporter is virtually turned
CC off at neutral pH;
CC CC -1- SUBUNIT: Interacts with tescalcin, CHP and CHP2.
CC CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC CC -1- TISSUE SPECIFICITY: Kidney and intestine.

CC CC -1- PTM: Phosphorylated (Possible).
CC CC -1- MISCELLANEOUS: Inhibited by amiloride and 5-amino-substituted
CC derivatives activated in a cooperative fashion by
CC intracellular H⁺. In quiescent cells upon growth factor
CC stimulation, the apparent affinity for internal H⁺ is increased,
CC resulting in a persistent rise in cytoplasmic pH.
CC CC -1- SIMILARITY: Belongs to the Na⁺/H⁺ exchanger family.
CC CC -1- CAUTION: The number, localization and denomination of hydrophobic
CC domains in the Na⁺/H⁺ exchangers vary among authors.
CC CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC CC
CC EMBL; M81768; AAB59460.1; ALT SEQ: mRNA.
CC EMBL; S68616; AAC60606.1; -; mRNA.
CC EMBL; AF141350; AAF21350.1; -; mRNA.
CC EMBL; AF141351; AAF21351.1; -; mRNA.
CC EMBL; AF141352; AAF21352.1; -; mRNA.
CC EMBL; AF141353; AAF21353.1; -; mRNA.
CC EMBL; AF141354; AAF21354.1; -; mRNA.
CC EMBL; AF141355; AAF21355.1; -; mRNA.
CC EMBL; AF141356; AAF21356.1; -; mRNA.
CC EMBL; AF141357; AAF21357.1; -; mRNA.
CC EMBL; AF141358; AAF21358.1; -; mRNA.
CC EMBL; AF141359; AAF21359.1; -; mRNA.
CC EMBL; AF146430; AAF25592.1; -; mRNA.
CC EMBL; AF146431; AAF25593.1; -; mRNA.
CC EMBL; AF146432; AAF25594.1; -; mRNA.
CC EMBL; AF146433; AAF25595.1; -; mRNA.
CC EMBL; AF146434; AAF25596.1; -; mRNA.
CC EMBL; AF146435; AAF25597.1; -; mRNA.
CC EMBL; AF146436; AAF25598.1; -; mRNA.
CC EMBL; AF146437; AAF25599.1; -; mRNA.
CC EMBL; AF146438; AAF25600.1; -; mRNA.
CC EMBL; AF146439; AAF25601.1; -; mRNA.
CC PIR; I57487; I57487.
CC PDB; 1V4E; NMR; A=155-180.
CC Ensembl; ENSG00000090020; Homo sapiens.
CC HGNC; HGNC:11071; SLC9A1.
CC MIM; 107310; -.
CC GO; GO:0016021; C: integral to membrane; TAS.
CC GO; GO:0015299; F: solute:hydrogen antiporter activity; TAS.
CC GO; GO:0006885; P: regulation of pH; TAS.
CC GO; GO:0006153; Na⁺ H⁺ porter.
CC InterPro; IPR004709; NaH_exchng.
CC InterPro; IPR001970; NaH_exchng.
CC Pfam; PF00999; Na_H_Exchange; 1.
CC PRINTS; PR01084; NAHEXCHNGR.
CC PRINTS; PR01085; NAHEXCHNGR1.
CC TIGRFAMs; TIGR00840; b_cpai; 1.
CC 3D-structure; Antiport; Glycoprotein; Ion transport; Multigene family;
CC Phosphorylation; Sodium; Sodium transport; Transmembrane; Transport.
FT TOPO_DOM 1 15 Cytoplasmic (Potential).
FT TRANSMEM 16 35 Extracellular (Potential).
FT TOPO_DOM 36 107 M2 (Potential).
FT TRANSMEM 108 127 M3 (Potential).
FT TOPO_DOM 128 129 M4 (Potential).
FT TRANSMEM 130 149 M5 (Potential).
FT TOPO_DOM 150 154 M6 (Potential).
FT TRANSMEM 155 174 M7 (Potential).
FT TOPO_DOM 175 191 M5 (Potential).
FT TRANSMEM 192 211 Extracellular (Potential).
FT TOPO_DOM 212 227 M5A (Potential).
FT TRANSMEM 228 247 M5B (Potential).
FT TOPO_DOM 248 256 M5C (Potential).
FT TRANSMEM 257 276 M5D (Potential).
FT TOPO_DOM 277 294 M6 (Potential).
FT TRANSMEM 295 315 M6 (Potential).
FT TOPO_DOM 316 338 M7 (Potential).
FT TRANSMEM 339 358 M7 (Potential).

FT TOPO_DOM 359 386 Extracellular (Potential).
 FT TRANSMEM 387 406 M8 (Potential).
 FT TOPO_DOM 407 410 Cytoplasmic (Potential).
 FT TRANSMEM 411 430 M9 (Potential).
 FT TOPO_DOM 431 480 Extracellular (Potential).
 FT TRANSMEM 481 500 M10 (Potential).
 FT TOPO_DOM 501 815 Cytoplasmic (Potential).
 FT CARBOHYD 75 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 370 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 815 AA; 90763 MW; 02EC748C79DF6526 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 815;
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLGL 89
 |||||
 Db 738 KGKVLGL 744

RESULT 761

QSVU07 HUMAN
 ID QSVU07 HUMAN PRELIMINARY; PRT; 815 AA.
 AC QSVU07;
 DT 01-FEB-2005 (T-EMBLrel. 29, Created)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
 DB Solute carrier family 9 (Sodium/hydrogen exchanger), isoform 1
 DE (Antipporter, Na+/H+, amiloride sensitive).
 GN Names=SLC9A1; ORFNames=RP4-633N17.1-001;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OC NCBI_TaxID=9606;
 RN NUCLEOTIDE SEQUENCE.
 RP Harrison E.;
 RA Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RL NUCLEOTIDE SEQUENCE.

RA Hall R.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL590640; CAI22090.1; -; Genomic DNA.
 DR EMBL; AL137860; CAI22090.1; -; Genomic DNA.
 DR EMBL; AL590640; CAI22090.1; JOINED; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0015385; F:solute:hydrogen antiporter activity; IEA.
 DR GO; GO:0015299; F:solute:hydrogen antiporter activity; IEA.
 DR GO; GO:0006885; P:regulation of pH; IEA.
 DR GO; GO:0006814; P:sodium ion transport; IEA.
 DR InterPro; IPR006153; Na H porter.
 DR InterPro; IPR004709; NaH exchanger.
 DR InterPro; IPR001970; NaH_exchng_r_1.
 DR Pfam; PF00999; Na_H_Exchange; 1.
 DR PRINTS; PR01084; NAHEXCHNGR.
 DR PRINTS; PR01085; NAHEXCHNGR1.
 DR TIGRFAMs; TIGR00840; b_cpai1.1.
 SQ SEQUENCE 815 AA; 90763 MW; 02EC748C79DF6526 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 815;
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLGL 89
 |||||
 Db 738 KGKVLGL 744

RESULT 762

SL9A1 RABIT
 ID SL9A1_RABIT STANDARD; PRT; 816 AA.

P23791;
 AC 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Sodium/hydrogen exchanger 1 (Na(+)/H(+)) exchanger 1 (NHE-1) (Solute carrier family 9 member 1).
 GN Names=SLC9A1; Synonyms=NHE1;
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
 OC Oryctolagus.
 OC NCBI_TaxID=9986;
 RN NUCLEOTIDE SEQUENCE.
 RP STRAIN=New Zealand white; TISSUE=ileal villus;
 RC MEDLINE=91293066; PubMed=1712287;
 RA Tee C.-M., Ma A.-I., Yang V.W., Watson A.J.M., Levine S.,
 RA Montrose M.H., Potter J., Sardet C., Pouyssegur J., Donowitz M.;
 RT "Molecular cloning and expression of a cDNA encoding the rabbit ileal villus cell basolateral membrane Na+/H+ exchanger.";
 RL EMBO J. 10:1957-1967(1991).
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Kidney;
 RC MEDLINE=92096447; PubMed=1661611; DOI=10.1016/0167-4781(91)90221-7;
 RA Hildebrandt F., Pizzonia J.H., Reilly R.F., Reboucas N.A., Sardet C.,
 RA Pouyssegur J., Slayman C.W., Aronson P.S., Igaraahi P.;
 RT "Cloning, sequence, and tissue distribution of a rabbit renal Na+/H+ exchanger transcript.";
 RL Blochim. Biophys. Acta 1129:105-108(1991).
 RN NUCLEOTIDE SEQUENCE OF 472-816.
 RP STRAIN=New Zealand white; TISSUE=Heart muscle;
 RC MEDLINE=91138752; PubMed=1704856; DOI=10.1016/0014-5793(91)80241-T;
 RA Fliegel L., Sardet C., Pouyssegur J., Barr A.;
 RT "Identification of the protein and cDNA of the cardiac Na+/H+ exchanger.";
 RL FEBS Lett. 279:25-29(1991).
 CC -/- FUNCTION: Involved in pH regulation to eliminate acids generated by active metabolism or to counter adverse environmental conditions. Major proton extruding system driven by the inward sodium ion chemical gradient. Plays an important role in signal transduction.
 CC -/- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -/- TISSUE SPECIFICITY: Kidney and intestine.
 CC -/- PTM: Phosphorylated (Possible).
 CC -/- SIMILARITY: Belongs to the Na(+)/H(+) exchanger family.
 CC -/- CAUTION: The number, localization and denomination of hydrophobic domains in the Na(+)/H(+) exchangers vary among authors.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
 CC -----
 DR EMBL; X59935; CAA42558.1; -; mRNA.
 DR EMBL; X61504; CAA43721.1; -; mRNA.
 DR EMBL; X56536; CAA39881.1; -; mRNA.
 DR PIR; S16328; S16328.
 DR InterPro; IPR006153; Na H porter.
 DR InterPro; IPR004709; NaH_exchang.
 DR InterPro; IPR001970; NaH_exchng_r_1.
 DR Pfam; PF00999; Na_H_Exchange; 1.
 DR PRINTS; PR01084; NAHEXCHNGR.
 DR PRINTS; PR01085; NAHEXCHNGR1.
 DR TIGRFAMs; TIGR00840; b_cpai1.1.
 DR Antipor; Glycoprotein; ion transport; Multigene family;
 KW Phosphorylation; Sodium; Sodium transport; Transmembrane; Transport.
 FT TOPO_DOM 1 15 Cytoplasmic (Potential).
 FT TRANSMEM 16 35 M1 (Potential).
 FT TOPO_DOM 36 107 Extracellular (Potential).
 FT TRANSMEM 108 127 M2 (Potential).

```

FT TOPO DOM 128 129 Cytoplasmic (Potential).
FT TRANS MEM 130 149 M3 (Potential).
FT TOPO DOM 150 154 Extracellular (Potential).
FT TRANS MEM 155 171 M4 (Potential).
FT TOPO DOM 175 191 Cytoplasmic (Potential).
FT TRANS MEM 192 211 M5 (Potential).
FT TOPO DOM 212 227 Extracellular (Potential).
FT TRANS MEM 228 247 M5A (Potential).
FT TOPO DOM 248 256 M5B (Potential).
FT TRANS MEM 257 276 Extracellular (Potential).
FT TOPO DOM 277 294 Extracellular (Potential).
FT TRANS MEM 295 315 M6 (Potential).
FT TOPO DOM 316 338 Cytoplasmic (Potential).
FT TRANS MEM 339 358 M7 (Potential).
FT TOPO DOM 359 386 Extracellular (Potential).
FT TRANS MEM 387 406 M8 (Potential).
FT TOPO DOM 407 410 M9 (Potential).
FT TRANS MEM 411 430 Extracellular (Potential).
FT TOPO DOM 431 480 M10 (Potential).
FT TRANS MEM 481 501 Cytoplasmic (Potential).
FT TOPO DOM 501 816 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 75 75 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 370 370 V -> A (in Ref. 2).
FT CONFLICT 242 242 K -> E (in Ref. 2).
FT CONFLICT 569 569
SQ SEQUENCE 816 AA; 90717 MW; 336738D267F7F436 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 816;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLGL 89
Db 738 KGVVLGL 744

RESULT 763
SL9A1_BOVIN STANDARD; PRT; 817 AA.
AC Q28036;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Sodium/hydrogen exchanger 1 (Na(+)/H(+)) (Solute
DE carrier family 9 member 1).
GN Name=SLC9A1; Synonyms=NHE1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RC TISSUE=Heart;
RA Zhu H., Zhang Q., Liu W., Trumbly R.J., Garlid K.D., Sun X.;
RT "Molecular cloning and characterization of Na+/H+ antiporter from
RT Bovine heart."
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Involved in pH regulation to eliminate acids generated
CC by active metabolism or to counter adverse environmental
CC conditions. Major proton extruding system driven by the inward
CC sodium ion chemical gradient. Plays an important role in signal
CC transduction.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: Phosphorylated (Potential).
CC -!- SIMILARITY: Belongs to the Na(+)/H(+) exchanger family.
CC -!- CAUTION: The number, localization and denomination of hydrophobic
CC domains in the Na(+)/H(+) exchangers vary among authors.
CC -!- CAUTION: Hydrophobic domains A, B and L are not believed to be
CC transmembranal, but only membrane-associated.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; U49432; AAA91483.1; -; mRNA.
DR InterPro; IPR006153; Na_H_porter.
DR InterPro; IPR004709; NaH_exchng.
DR InterPro; IPR001970; NaH_exchng.1.
DR Pfam; PF00999; Na_H_Exchange; 1.
DR PRINTS; PR01084; NAHEXCHNGR.
DR PRINTS; PR01085; NAHEXCHNGR1.
DR TIGRPFAMs; TIGR00840; b_cpai; 1.
KW Antiport; Glycoprotein; Ion transport; Multigene family;
KW Phosphorylation; Sodium; Sodium transport; Transmembrane; Transport.
FT TRANS MEM 1 15 Cytoplasmic (Potential).
FT TOPO DOM 16 35 M1 (Potential).
FT TRANS MEM 36 107 Extracellular (Potential).
FT TOPO DOM 108 127 M2 (Potential).
FT TRANS MEM 128 129 Cytoplasmic (Potential).
FT TOPO DOM 130 149 M3 (Potential).
FT TRANS MEM 150 154 Extracellular (Potential).
FT TOPO DOM 155 174 M4 (Potential).
FT TRANS MEM 175 191 Cytoplasmic (Potential).
FT TOPO DOM 192 211 M5 (Potential).
FT TOPO DOM 212 227 Extracellular (Potential).
FT TRANS MEM 228 247 M5A (Potential).
FT TOPO DOM 248 256 M5B (Potential).
FT TRANS MEM 257 276 Extracellular (Potential).
FT TOPO DOM 277 294 M6 (Potential).
FT TRANS MEM 295 315 M7 (Potential).
FT TOPO DOM 316 338 M7 (Potential).
FT TRANS MEM 339 358 Extracellular (Potential).
FT TOPO DOM 359 386 M8 (Potential).
FT TRANS MEM 387 406 Cytoplasmic (Potential).
FT TOPO DOM 407 410 M9 (Potential).
FT TRANS MEM 411 430 Extracellular (Potential).
FT TOPO DOM 431 480 M10 (Potential).
FT TRANS MEM 481 501 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 75 75 N-linked (GlcNAc. .) (Potential).
FT CARBOHYD 370 370 N-linked (GlcNAc. .) (Potential).
SQ SEQUENCE 817 AA; 91017 MW; 6617E99D3B012920 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 817;
Best Local Similarity 100.0%; Pred. No. 8.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLGL 89
Db 738 KGVVLGL 744

RESULT 764
SL9A1_PIG STANDARD; PRT; 818 AA.
AC P48762;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Sodium/hydrogen exchanger 1 (Na(+)/H(+)) (Solute
DE carrier family 9 member 1).
GN Name=SLC9A1; Synonyms=NHE1;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
OC Sus.
OX NCBI_TaxID=9823;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92087905; PubMed=1661081;
RA Reilly R.F., Hildebrandt F., Biemesderfer D., Sardet C.,
RA Pouyssegur J., Aronson P.S., Slayman C.W., Igazashi P.;
RT "cDNA cloning and immunolocalization of a Na(+)-H+ exchanger in LLC-

```

RT PK1 renal epithelial cells.";
 RL Am. J. Physiol. 261:F1088-F1094 (1991).
 CC -1- FUNCTION: Involved in pH regulation to eliminate acids generated
 CC by active metabolism or to counter adverse environmental
 CC conditions. Major proton extruding system driven by the inward
 CC sodium ion chemical gradient. Plays an important role in signal
 CC transduction.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: Phosphorylated (possible).
 CC -1- SIMILARITY: Belongs to the Na(+)/H(+) exchanger family.
 CC -1- CAUTION: The number, localization and denomination of hydrophobic
 CC domains in the Na(+)/H(+) exchangers vary among authors.
 CC -1- CAUTION: Hydrophobic domains A, B and L are not believed to be
 CC transmembranal, but only membrane-associated.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; M99631; AAA31092.1; -; mRNA.
 DR EMBL; S71135; AAB20633.1; -; mRNA.
 DR PIR; A48858.
 DR InterPro; IPR006153; Na_H_porter.
 DR InterPro; IPR004709; NaH_exchang.
 DR InterPro; IPR001970; NaH_exchangr_1.
 DR Pfam; PF00999; Na_H_Exchange; 1.
 DR PRINTS; PR01084; NAHEXCHNGR.
 DR PRINTS; PR01085; NAHEXCHNGR.
 DR TIGRfams; TIGR00840; b_cpai; 1.
 KW Antiport; Glycoprotein; Ion transport; Multigene family;
 KW Phosphorylation; Sodium; Sodium transport; Transmembrane; Transport.
 FT TOPO_DOM 1 11
 FT CYTOPLASMIC (Potential).
 FT TOPO_DOM 32 101
 FT TOPO_DOM 124 126
 FT TRANSMEM 127 146
 FT TRANSMEM 147 158
 FT TOPO_DOM 147 158
 FT TRANSMEM 159 179
 FT TOPO_DOM 180 184
 FT TRANSMEM 185 206
 FT TOPO_DOM 207 226
 FT TRANSMEM 227 247
 FT TOPO_DOM 248 256
 FT TRANSMEM 257 278
 FT TOPO_DOM 279 297
 FT TRANSMEM 298 318
 FT TOPO_DOM 319 332
 FT TRANSMEM 333 353
 FT TOPO_DOM 354 384
 FT TRANSMEM 385 406
 FT TOPO_DOM 407 412
 FT TRANSMEM 413 434
 FT TOPO_DOM 435 448
 FT TOPO_DOM 470 478
 FT TRANSMEM 479 499
 FT TOPO_DOM 500 818
 FT REGION 12 31
 FT REGION 102 123
 FT REGION 449 469
 FT CARBOHYD 370 370
 FT CONFLICT 683 683
 SQ SEQUENCE 818 AA; 90987 MW; 932979DA51D3DC9 CRC64;
 Query Match 5.9%; Score 7; DB 1; Length 818;
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 83 KGKVLGL 89
 Db 738 KGKVLGL 744

RESULT 765
 SL9A1 MOUSE STANDARD; PRT; 820 AA.
 ID SL9A1_MOUSE Q61165;
 AC Q61165;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DT Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1) (Solute
 DR carrier family 9 member 1).
 DR Name=Slc9a1; Synonyms=Nhel;
 DR Mus musculus (Mouse).
 DR Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 DR Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 DR Muroidae; Muridae; Murinae; Mus.
 DR NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=BALB/c;
 RX MEDLINE=21928362; PubMed=11931388; DOI=10.1023/A:1017984311138;
 RA Dewey M.J., Emnis T.M., Bowman L.H.;
 RT "cDNA cloning and expression of the mouse Na/H antiporter (NHE-1) and
 RL Mol. Biol. Rep. 28:111-117 (2001).
 CC -1- FUNCTION: Involved in pH regulation to eliminate acids generated
 CC by active metabolism or to counter adverse environmental
 CC conditions. Major proton extruding system driven by the inward
 CC sodium ion chemical gradient. Plays an important role in signal
 CC transduction.
 CC -1- SUBUNIT: Interacts with tescalcin, CHP and CHP2 (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: Phosphorylated (possible).
 CC -1- SIMILARITY: Belongs to the Na(+)/H(+) exchanger family.
 CC -1- CAUTION: The number, localization and denomination of hydrophobic
 CC domains in the Na(+)/H(+) exchangers vary among authors.
 CC -1- CAUTION: Hydrophobic domains A, B and L are not believed to be
 CC transmembranal, but only membrane-associated.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR EMBL; U51112; AAA92976.1; -; mRNA.
 DR EMBL; ENSMUSG00000028854; Mus musculus.
 DR MGI; MGI:102462; Slc9a1.
 DR GO; GO:0016323; C:basolateral plasma membrane; TAS.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR GO; GO:0005624; C:membrane fraction; IDA.
 DR GO; GO:0015385; F:sodium:hydrogen antiporter activity; IDA.
 DR GO; GO:0006885; P:regulation of pH; IDA.
 DR GO; GO:0006814; P:sodium ion transport; IDA.
 DR InterPro; IPR006153; Na_H_porter.
 DR InterPro; IPR004709; NaH_exchang.
 DR InterPro; IPR001970; NaH_exchangr_1.
 DR Pfam; PF00999; Na_H_Exchange; 1.
 DR PRINTS; PR01084; NAHEXCHNGR.
 DR PRINTS; PR01085; NAHEXCHNGR.
 DR TIGRfams; TIGR00840; b_cpai; 1.
 KW Antiport; Glycoprotein; Ion transport; Multigene family;
 KW Phosphorylation; Sodium; Sodium transport; Transmembrane; Transport.
 FT TOPO_DOM 1 12
 FT TOPO_DOM 33 105
 FT TOPO_DOM 128 130
 FT TRANSMEM 131 150
 FT TRANSMEM 151 162
 FT TOPO_DOM 151 162
 FT TRANSMEM 163 183
 FT TOPO_DOM 184 188
 FT TRANSMEM 189 210
 FT TOPO_DOM 211 230
 FT TRANSMEM 231 251
 FT TOPO_DOM 252 260
 FT CYTOPLASMIC (Potential).
 FT CYTOPLASMIC (Potential).
 FT C (M3) (Potential).
 FT Extracellular (Potential).
 FT D (M4) (Potential).
 FT E (M5) (Potential).
 FT Extracellular (Potential).
 FT F (M5A) (Potential).
 FT Cytoplasmic (Potential).

FT TRANSMEM 261 282 G (M5B) (Potential).
 FT TOPO DOM 283 301 Extracellular (Potential).
 FT TRANSMEM 302 322 H (M6) (Potential).
 FT TOPO DOM 323 336 I (M7) (Potential).
 FT TRANSMEM 337 357 Extracellular (Potential).
 FT TOPO DOM 358 388 J (M8) (Potential).
 FT TRANSMEM 389 410 Cytoplasmic (Potential).
 FT TOPO DOM 411 416 K (M9) (Potential).
 FT TRANSMEM 417 438 Extracellular (Potential).
 FT TOPO DOM 439 452 M (M10) (Potential).
 FT TRANSMEM 453 503 Cytoplasmic (Potential).
 FT TOPO DOM 504 820 A (M1) hydrophobic.
 FT REGION 13 32 B (M2) hydrophobic.
 FT REGION 106 127 L, hydrophobic.
 FT CARBOHYD 453 473 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 374 374 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 820 AA; 91468 MW; 0589C4D08DD348BE CRC64;

Query Match 5.9%; Score 7; DB 1; Length 820;
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLGL 89
 Db 742 KGVVLGL 748

RESULT 766
 SL9A1 RAT STANDARD; PRT; 820 AA.
 AC P26431;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Sodium/hydrogen exchanger 1 (Na(+)/H(+) exchanger 1) (NHE-1) (Solute carrier family 9 member 1).
 GN Rattus norvegicus (Rat).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Murinae; Murinae; Rattus.
 OK NCBI_TaxID=10116;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Sprague-Dawley; TISSUE=Heart;
 RX MEDLINE=92250539; PubMed=157762;
 RA Orłowski J., Kandasamy R.A., Shull G.E.;
 RT "Molecular cloning of putative members of the Na/H exchanger gene family. cDNA cloning, deduced amino acid sequence, and mRNA tissue expression of the rat Na/H exchanger NHE-1 and two structurally related proteins.";
 RL J. Biol. Chem. 267:9331-9339(1992).
 CC -!- FUNCTION: Involved in pH regulation to eliminate acids generated by active metabolism or to counter adverse environmental conditions. Major proton extruding system driven by the inward sodium ion chemical gradient. Plays an important role in signal transduction.
 CC -!- SUBUNIT: Interacts with tescalcin, CHP and CHP2 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Not tissue specific.
 CC -!- PTM: Phosphorylated (Possible).
 CC -!- SIMILARITY: Belongs to the Na(+)/H(+) exchanger family.
 CC -!- CAUTION: The number, localization and denomination of hydrophobic domains in the Na(+)/H(+) exchangers vary among authors.
 CC -!- CAUTION: Hydrophobic domains A, B and L are not believed to be transmembranal, but only membrane-associated.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

CC EMBL; M85299; AAA98479.1; -; mRNA.
 DR Ensembl; ENSRNOG0000007982; Rattus norvegicus.
 DR RGD; 3718; Slc9a1.
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0015385; F:sodium:hydrogen antiporter activity; TAS.
 DR GO; GO:0008134; F:transcription factor binding; IDA.
 DR GO; GO:0006885; P:regulation of pH; TAS.
 DR GO; GO:0006814; P:sodium ion transport; TAS.
 DR InterPro; IPR006153; Na H porter.
 DR InterPro; IPR004709; NaH_exchang.
 DR InterPro; IPR001970; NaH_exchng_1.
 DR Pfam; PF00999; Na H Exchanger; 1.
 DR PRINTS; PR01084; NAHEXCHNGR1.
 DR TIGRFAMs; TIGR00840; b_cpai; 1.
 KW Anticpport; Glycoprotein; Ion transport; Multigene family; Transport.
 KW Phosphorylation; Sodium; Sodium transport; Transmembrane; Transport.
 FT TOPO DOM 1 12 Cytoplasmic (Potential).
 FT TOPO DOM 33 105 Cytoplasmic (Potential).
 FT TOPO DOM 128 130 Cytoplasmic (Potential).
 FT TRANSMEM 131 150 C (M3) (Potential).
 FT TOPO DOM 151 162 Extracellular (Potential).
 FT TRANSMEM 163 183 D (M4) (Potential).
 FT TOPO DOM 184 188 Cytoplasmic (Potential).
 FT TRANSMEM 189 210 E (M5) (Potential).
 FT TOPO DOM 211 230 Extracellular (Potential).
 FT TRANSMEM 231 251 F (M5A) (Potential).
 FT TOPO DOM 252 260 Cytoplasmic (Potential).
 FT TRANSMEM 261 282 G (M5B) (Potential).
 FT TOPO DOM 283 301 Extracellular (Potential).
 FT TRANSMEM 302 322 H (M6) (Potential).
 FT TOPO DOM 323 336 I (M7) (Potential).
 FT TRANSMEM 337 357 Extracellular (Potential).
 FT TOPO DOM 358 388 J (M8) (Potential).
 FT TRANSMEM 389 410 Cytoplasmic (Potential).
 FT TOPO DOM 411 416 K (M9) (Potential).
 FT TRANSMEM 417 438 Extracellular (Potential).
 FT TOPO DOM 439 452 Extracellular (Potential).
 FT TOPO DOM 474 482 Extracellular (Potential).
 FT TRANSMEM 483 503 M (M10) (Potential).
 FT TOPO DOM 504 820 Cytoplasmic (Potential).
 FT REGION 13 32 A (M1) hydrophobic.
 FT REGION 106 127 B (M2) hydrophobic.
 FT CARBOHYD 453 473 L, hydrophobic.
 FT CARBOHYD 374 374 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 820 AA; 91647 MW; 58398DE74A9642FB CRC64;

Query Match 5.9%; Score 7; DB 1; Length 820;
 Best Local Similarity 100.0%; Pred. No. 8.8e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGVVLGL 89
 Db 742 KGVVLGL 748

RESULT 767
 Q80X31_MOUSE
 ID Q80X31_MOUSE PRELIMINARY; PRT; 820 AA.
 AC Q80X31;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Slc9a1 protein.
 GN Name=Slc9a1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; OC Muridae; Murinae; Mus.
 OK NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.

EMBL; AE016944; AAO79442.1; -; Genomic_DNA.
HSSP; P07686; INOW.
GO; GO:0004563; F:beta-N-acetylhexosaminidase activity; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001540; Glyco_hydro_20.
Pfam; PF00728; Glyco_hydro_20; 1.
PRINTS; PR00738; GLYDRLASE20.
KW Complete proteome.
SQ SEQUENCE 844 AA; 94937 MW; E71BEC4DFB56C471 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 844;
Best Local Similarity 100.0%; Pred. No. 9e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 LAALAAAY 27
Db 315 LAALAAAY 321
RESULT 770
ID Q8X316 ECO57
QX Q8X316 ECO57 PRELIMINARY; PRT; 870 AA.
AC Q8X316;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative tail length tape measure protein.
GN OrderedLocNames=ECs2166;
OS Escherichia coli O157:H7;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Toke T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RL EMBL; BA000007; BAB35589.1; -; Genomic_DNA.
DR PIR; F90899; F90899.
DR InterPro; IPR006431; Tape_meas_lam_C.
DR InterPro; IPR009628; TMP_2.
DR Pfam; PF06791; TMP_2; 1.
DR TIGRFAMs; TIGR01541; tape_meas_lam_C; 1.
SQ SEQUENCE 870 AA; 93456 MW; 6BC7DA2A6A1EB5EC CRC64;
Query Match 5.9%; Score 7; DB 2; Length 870;
Best Local Similarity 100.0%; Pred. No. 9.2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALAA 26
Db 166 VLAALAA 172
RESULT 771
ID Q8TV7 TRYCR
QX Q8TV7 TRYCR PRELIMINARY; PRT; 875 AA.
AC Q8TV7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proton motive P-type ATPase 1.
GN Name=HAL;
OS Trypanosoma cruzi.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN NUCLEOTIDE SEQUENCE.
RP Creutzburg K.; Koehler B., Hempel H., Schreier P. Jacobs E., Jacobs E.,
RA Schmidt H.;
RT "Genetic structure and chromosomal integration site of the cryptic
RT Shiga toxin 1-converting prophage CP-1639.";
RL Microbiol. 151:941-950(2005).
RN NUCLEOTIDE SEQUENCE.
RP Creutzburg K.;
RA Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ304858; CAH23254.1; -; Genomic DNA.
DR

EMBL; AE016944; AAO79442.1; -; Genomic_DNA.
HSSP; P07686; INOW.
GO; GO:0004563; F:beta-N-acetylhexosaminidase activity; IEA.
GO; GO:0005975; P:carbohydrate metabolism; IEA.
InterPro; IPR001540; Glyco_hydro_20.
Pfam; PF00728; Glyco_hydro_20; 1.
PRINTS; PR00738; GLYDRLASE20.
KW Complete proteome.
SQ SEQUENCE 844 AA; 94937 MW; E71BEC4DFB56C471 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 844;
Best Local Similarity 100.0%; Pred.No. 9e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 LAALAAAY 27
Db 315 LAALAAAY 321
RESULT 770
ID Q8X316 ECO57
QX Q8X316 ECO57 PRELIMINARY; PRT; 870 AA.
AC Q8X316;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative tail length tape measure protein.
GN OrderedLocNames=ECs2166;
OS Escherichia coli O157:H7;
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=O157:H7 / Sakai / RIMD 050952 / EHEC;
RC MEDLINE=21156231; PubMed=11258796;
RA Hayaishi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RL EMBL; BA000007; BAB35589.1; -; Genomic_DNA.
DR PIR; F90899; F90899.
DR InterPro; IPR006431; Tape_meas_lam_C.
DR InterPro; IPR009628; TMP_2.
DR Pfam; PF06791; TMP_2; 1.
DR TIGRFAMs; TIGR01541; tape_meas_lam_C; 1.
SQ SEQUENCE 870 AA; 93456 MW; 6BC7DA2A6A1EB5EC CRC64;
Query Match 5.9%; Score 7; DB 2; Length 870;
Best Local Similarity 100.0%; Pred.No. 9.2e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALAA 26
Db 166 VLAALAA 172
RESULT 771
ID Q8TV7 TRYCR
QX Q8TV7 TRYCR PRELIMINARY; PRT; 875 AA.
AC Q8TV7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Proton motive P-type ATPase 1.
GN Name=HAL;
OS Trypanosoma cruzi.
OC Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN NUCLEOTIDE SEQUENCE.
RP Creutzburg K.;
RA Creutzburg K.;
RA Schmidt H.;
RT "Genetic structure and chromosomal integration site of the cryptic
RT Shiga toxin 1-converting prophage CP-1639.";
RL Microbiol. 151:941-950(2005).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Creutzburg K.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ304858; CAH23254.1; -; Genomic DNA.

```

DR InterPro; IPR006431; Tape_meas_lam_C.
DR InterPro; IPR009628; TWP_2.
DR Pfam; PF06791; TWP_2; 1.
DR TIGRFAMs; TIGR01541; tape_meas_lam_C; 1.
SQ SEQUENCE 881 AA; 94431 MW; 8AB4B708A59FA1F0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 881;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 177 VLAALAA 183

RESULT 773
ID Q9EYEL_ECO57 PRELIMINARY; PRT; 881 AA.
AC Q9EYEL
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Minor tail protein H (putative tail length tape measure protein).
GN OrderedLocuNames=ECs2949;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1111050; DOI=10.1016/S0378-1119(00)00416-9;
RA Yokoyama K., Makino K., Kubota Y., Watanabe M., Kimura S.,
RA Yutsudo C.H., Kurokawa K., Ishii K., Hattori M., Tatsuno I., Abe H.,
RA Itoh M., Iida T., Ohnishi M., Hayashi T., Yasunaga T., Honda T.,
RA Sasaki K., Shinagawa H.;
RT "Complete nucleotide sequence of the prophage VT1-Sakai carrying the
RT Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7
RT strain derived from the Sakai outbreak.";
RL Gene 258:127-139(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=O157:H7 / Sakai / RMD 0509952 / EHRC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Hattori M., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001)
DR EMBL; AP000400; BAB19568.1; -; Genomic_DNA.
DR EMBL; BA000007; BAB36372.1; -; Genomic_DNA.
DR FIRM; E90997; E90997.
DR InterPro; IPR006431; Tape_meas_lam_C.
DR InterPro; IPR009628; TWP_2.
DR Pfam; PF06791; TWP_2; 1.
DR TIGRFAMs; TIGR01541; tape_meas_lam_C; 1.
SQ SEQUENCE 881 AA; 94562 MW; 99132D29E6A1A6E2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 881;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 177 VLAALAA 183

RESULT 774
ID Q9I537_PSEAE PRELIMINARY; PRT; 881 AA.
AC Q9I537
DT 01-MAR-2001 (TrEMBLrel. 16, Created)

```

```

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocuNames=PA0920;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=2043737; PubMed=10984043; DOI=10.1038/35023079;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Lim R.M.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H. Jr., Hancock R.E.W., Lozy S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AE004526; AAG04309.1; -; Genomic_DNA.
DR FIRM; F83530; F83530.
DR InterPro; IPR007424; DUF470.
DR InterPro; IPR007425; DUF471.
DR InterPro; IPR007426; DUF472.
DR Pfam; PF04329; DUF470; 1.
DR Pfam; PF04330; DUF471; 1.
DR Pfam; PF04331; DUF472; 1.
DR Complete proteome; Hypothetical protein.
KW SEQUENCE 881 AA; 95861 MW; F9C595E70BAA0C57 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 881;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 166 VLAALAA 172

RESULT 775
Q8X327_ECO57 PRELIMINARY; PRT; 884 AA.
ID Q8X327
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative tail fiber component H of prophage CP-933U.
GN OrderedLocuNames=z3084;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHRC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
DR EMBL; AE005174; AAG57000.1; -; Genomic_DNA.
DR FIRM; D85817; D85817.
DR InterPro; IPR006431; Tape_meas_lam_C.
DR InterPro; IPR009628; TWP_2.
DR Pfam; PF06791; TWP_2; 1.
DR TIGRFAMs; TIGR01541; tape_meas_lam_C; 1.
KW Complete proteome.

```

```
SQ SEQUENCE 884 AA; 94837 MW; 72BE7C25AAEA9F1 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 884;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 180 VLAALAA 186

RESULT 776
Q7QET0 ANOQA
ID Q7QET0 ANOQA PRELIMINARY; PRT; 900 AA.
AC Q7QET0;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE ENSANGP0000019891 (Fragment).
GN ORFNames=ENSANGG0000017402;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=PEST;
RC The Anopheles gambiae Sequence Committee;
RG The Anopheles gambiae re-annotation";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AA0801008846; EAA06388.2; -; Genomic_DNA.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR_1; 7.
DR SMART; SM00028; TPR_12.
DR PROSITE; PS00005; TPR; 9.
DR PROSITE; PS0293; TPR_REGION; 2.
KW TPR repeat.
FT NON TER
SQ SEQUENCE 900 AA; 102321 MW; B9C86C0BC9B5036D CRC64;
Query Match 5.9%; Score 7; DB 2; Length 900;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAALA 25
Db 504 GVLAALA 510

RESULT 777
Q53964 STRCO
ID Q53964 STRCO PRELIMINARY; PRT; 914 AA.
AC Q53964;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Regulatory protein (fragment).
GN Name=whiB;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
```

```
RC STRAIN=A3;
RX MEDLINE=92269753; PubMed=1316997; DOI=10.1007/BF00266237;
RA Davis N.K., Chater K.F.;
RT "The Streptomyces coelicolor whiB gene encodes a small transcription
RT factor-like protein dispensable for growth but essential for
RT sporulation.";
RL Mol. Gen. Genet. 232:351-358(1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=A3;
RA Bruton C.J.;
RL Submitted (SEP-1991) to the EMBL/GenBank/DBJ databases.
DR EMBL; X62287; CAA44176.1; -; Genomic_DNA.
DR PIR; S18942; S18942.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF00535; Glycos transf_2; 1.
FT NON TER 914
SQ SEQUENCE 914 AA; 96022 MW; C6AB3B040698B43D CRC64;
Query Match 5.9%; Score 7; DB 2; Length 914;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
Db 719 VLAALAA 725

RESULT 778
Q8T7V6 TRYCR
ID Q8T7V6 TRYCR PRELIMINARY; PRT; 917 AA.
AC Q8T7V6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Proton motive P-type ATPase 2.
GN Name=HA2;
OS Trypanosoma cruzi.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22313545; PubMed=12221074; DOI=10.1074/jbc.M202267200;
RA Luo S., Scott D.A., Docampo R.;
RT "Trypanosoma cruzi H+-ATPase 1 (TcHA1) and 2 (TcHA2) genes complement
RT yeast mutants defective in H+ pumps and encode plasma membrane P-type
RT H+-ATPases with different enzymatic properties.";
RL J. Biol. Chem. 277:44497-44505(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; AF54412; AAL87542.1; -; Genomic_DNA.
DR HSSP; P07038; IMHS.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016820; F:hydrolase activity; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR006534; ATPase-IIIa H.
DR InterPro; IPR001757; ATPase E1-E2.
DR InterPro; IPR004014; Cation ATPase N.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR Pfam; PF00690; Cation ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
```

DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMS; TIGR01647; ATPase-IIIa_H; 1.
DR TIGRFAMS; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN 1.
KW ATP-binding; Hydrolase; Nucleotide-binding; Transmembrane.
SQ SEQUENCE 917 AA; 101255 MW; 37B23E7EF14F0774 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 917;
Best Local Similarity 100.0%; Pred. No. 9.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 377 VLAALAA 383
|||||

RESULT 779
O15637_TRYCR PRELIMINARY; PRT; 925 AA.
AC O15637;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Proton motive ATPase.
DE Name=TCB3;
OS Trypanosoma cruzi.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC Schizotrypanum.
OX NCBI_TaxID=5693;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21062729; PubMed=11077265; DOI=10.1016/S1383-5769(00)00061-1;
RA Meade J.C., Li C., Stiles J.K., Moate M.E., Penny J.I., Krishna S.,
RA Finley R.W.;
RT "The Trypanosoma cruzi genome contains ion motive ATPase genes which
RT closely resemble Leishmania proton pumps."
RL Parasitol. Int. 49:309-320(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
EMBL; AF000161; AAB70152.1; -; Genomic_DNA.
DR HSSP; P07038; IMHS
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0016887; F:ATPase activity; IEA.
DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. . .; IEA.
DR GO; GO:0006812; P:cation transport; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR006534; ATPase-IIIa_H.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR004014; Cation ATPase N.
DR InterPro; IPR005834; Dehal like hydro.
DR InterPro; IPR008250; E1-E2 ATPase_reg.
DR InterPro; IPR000695; H ATPase.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMS; TIGR01647; ATPase-IIIa_H; 1.
DR TIGRFAMS; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPase_E1_E2; UNKNOWN 1.
KW ATP-binding; Hydrolase; Nucleotide-binding; Transmembrane.
SQ SEQUENCE 925 AA; 102027 MW; 59DA91136489CD31 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 925;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
Db 377 VLAALAA 383
|||||

RESULT 780
Q8TBM9_HUMAN PRELIMINARY; PRT; 937 AA.
AC Q8TBM9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein MGC22014.
GN Name=MGC22014;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haiech P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX TISSUE=Duodenum;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Kozłowicz A., Doebber A., Goyea E.;
RT "The sequence of Homo sapiens BAC clone RP11-287D1.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCES.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022443; AAH22443.1; -; mRNA.
DR EMBL; AC073263; AAY93057.1; -; Genomic DNA.
DR Ensembl; ENSG00000187605; Homo sapiens.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001367; HTH_DtXR.

KW Hypothetical protein.
SQ SEQUENCE 937 AA; 102442 MW; 2B97D8348DB5F8B2 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 937;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 PAIVPDK 52
Db 607 PAIVPDK 613

RESULT 781
Q7VN93 HAE DU PRELIMINARY; PRT; 961 AA.
AC Q7VN93;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Valyl-CRNA synthetase.
GN Name=vals; OrderedLocusNames=HD0669;
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_taxid=730;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;
RA "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017152; AAP95593.1; -; Genomic_DNA.
DR HSSP; P96142; 11VS.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004832; P:valine-tRNA ligase activity; IEA.
DR GO; GO:0006438; P:valyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002300; trna-synt 1a.
DR InterPro; IPR001412; trna-synt 1.
DR InterPro; IPR002303; trna-synt_val.
DR Pfam; PF00133; trna-synt 1; 1.
DR PRINTS; PR00986; TRNASYNTHAL.
DR TIGRFAMs; TIGR00422; vals; 1.
DR PROSITE; PS00178; AA TRNA LIGASE I; 1.
KW Aminoacyl-tRNA synthetase; Complete proteome.
SQ SEQUENCE 961 AA; 108974 MW; EFDFA4C41C9238B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 961;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 94 TQQQAVI 100
Db 840 TQQQAVI 846

RESULT 782
ATXA LEIDO STANDARD; PRT; 974 AA.
AC P11718;
DT 01-OCT-1989 (Rel. 12, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Potential proton ATPase 1A (EC 3.6.3.6) (LDH1A protein).
GN Name=H1A;
OS Leishmania donovani.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX NCBI_taxid=5661;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=88122116; PubMed=2828921;
RA Meade J.C., Shaw J., Lenaster S., Gallagher G., Stringer J.R.;
RT "Structure and expression of a tandem gene pair in Leishmania donovani

RT that encodes a protein structurally homologous to eucaryotic cation-
RL transporting ATPases.";
RN Mol. Cell. Biol. 7:3937-3946(1987).
[2]
RP SEQUENCE REVISION TO 55-56, AND DEVELOPMENTAL STAGE.
RC STRAIN=MHOM/ET/67/L82;
RX MEDLINE=89219149; PubMed=2469011; DOI=10.1016/0166-6851(89)90045-5;
RA Meade J.C., Hudson K.M., Stringer S.L., Stringer J.R.;
RT "A tandem pair of Leishmania donovani cation transporting ATPase genes
RL encode isoforms that are differentially expressed.";
RN Mol. Biochem. Parasitol. 33:81-91(1989).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: More abundant in promastigotes than
CC amastigotes.
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family. Type IIIA subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; AF109296; AAA29227.2; -; Genomic_DNA.
DR HSSP; P07038; 1MHS.
DR InterPro; IPR006534; ATPase-IIIa H.
DR InterPro; IPR001757; ATPase_EI-E2.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehal_like_Hydro.
DR InterPro; IPR008250; EI-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR PANTHER; PTHR11939; ATPase_EI-E2; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; EI-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01647; ATPase-IIIa H; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPase_EI-E2; 1.
KW ATP-binding; Hydrogen ion transport; Hydrolase; Ion transport;
KW Magnesium; Metal-binding; Multigene family; Nucleotide-binding;
KW Phosphorylation; Transmembrane; Transport.
FT TOPO_DOM 1 92 Cytoplasmic (Potential).
FT TRANSMEM 93 112 1 (Potential).
FT TOPO_DOM 113 117 Extracellular (Potential).
FT TRANSMEM 118 137 2 (Potential).
FT TOPO_DOM 138 264 Cytoplasmic (Potential).
FT TRANSMEM 265 286 3 (Potential).
FT TOPO_DOM 287 294 Extracellular (Potential).
FT TRANSMEM 295 321 4 (Potential).
FT TOPO_DOM 322 630 Cytoplasmic (Potential).
FT TRANSMEM 631 651 5 (Potential).
FT TOPO_DOM 652 661 Extracellular (Potential).
FT TRANSMEM 662 684 6 (Potential).
FT TOPO_DOM 685 697 Cytoplasmic (Potential).
FT TRANSMEM 698 712 7 (Potential).
FT TOPO_DOM 713 737 Extracellular (Potential).
FT TRANSMEM 738 761 8 (Potential).
FT TOPO_DOM 762 812 Cytoplasmic (Potential).
FT TRANSMEM 813 840 9 (Potential).
FT TOPO_DOM 841 868 Extracellular (Potential).
FT TRANSMEM 869 887 10 (Potential).
FT TOPO_DOM 888 974 Cytoplasmic (Potential).
FT ACT_SITE 351 351 4-aspartylphosphate intermediate (By
FT similarity).
FT METAL 605 605 Magnesium (By similarity).
FT METAL 609 609 Magnesium (By similarity).
FT METAL 714 714 Magnesium (Potential).
FT MOD_RES 353 353 Phosphothreonine.
SQ SEQUENCE 974 AA; 107479 MW; 6BAB38F5153928A0 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 974;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 381 VLAALAA 387

RESULT 783
ATXB LEIDO STANDARD; PRT; 974 AA.

AC P12522; Rel. 12, Created
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Potential proton ATPase 1B (EC 3.6.3.6) (LDH1B protein).
GN Name-H1B;
OS Leishmania donovani.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5661;
RN NUCLEOTIDE SEQUENCE, AND DEVELOPMENTAL STAGE.
RC STRAIN=HOM/ET/67/L82;
RX MEDLINE=89219149; PubMed=2469011; DOI=10.1016/0166-6851(89)90045-5;
RT "A tandem pair of Leishmania donovani cation transporting ATPase genes
RL encode isoforms that are differentially expressed.";
RM Mol. Biochem. Parasitol. 33:81-91(1989).
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DEVELOPMENTAL STAGE: More abundant in amastigotes than
CC promastigotes.
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family. Type IIIA subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

EMBL; AF109296; AAA29228.1; -; Genomic_DNA.
DR HSP; P07038; IMHS.
DR InterPro; IPR006534; ATPase-III A H.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR PANTHER; PTHR11939; ATPase_E1-E2; 1.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01647; ATPase-III A H; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPase_E1-E2; 1.
KW ATP-binding; Hydrogen ion transport; Hydrolase; Ion transport;
KW Magnesium; Multigene family; Nucleotide-binding; Phosphorylation;
KW Transmembrane; Transport.

FT TRANSMEM 93 112 Potential.
FT TRANSMEM 118 137 Potential.
FT TRANSMEM 265 286 Potential.
FT TRANSMEM 295 321 Potential.
FT TRANSMEM 631 651 Potential.
FT TRANSMEM 662 684 Potential.
FT TRANSMEM 698 712 Potential.
FT TRANSMEM 738 761 Potential.
FT TRANSMEM 813 840 Potential.

FT TRANSMEM 869 887 Potential.
FT ACT_SITE 351 351 4-aspartylphosphate intermediate.
SQ SEQUENCE 974 AA; 107305 MW; 9F58AB61186556B1 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 974;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 381 VLAALAA 387

RESULT 784
Q4QDN8 LEIMA PRELIMINARY; PRT; 974 AA.

ID Q4QDN8 LEIMA PRELIMINARY;
AC Q4QDN8;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cation-transporting ATPase.
GN ORFNames=LmjF18.1510;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OC NCBI_TaxID=5664;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neil S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrall B.G.;
RA Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RL -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.

EMBL; CT005257; CAJ04293.1; -; Genomic_DNA.
DR InterPro; IPR006534; ATPase-III A H.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR004014; Cation_ATPase_N.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR Pfam; PF00690; Cation_ATPase_N; 1.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01647; ATPase-III A H; 1.
DR TIGRFAMs; TIGR01494; ATPase_P-type; 3.
DR PROSITE; PS00154; ATPase_E1-E2; UNKNOWN 1.
KW ATP-binding; Hydrolase; Nucleotide-binding; Transmembrane.
SQ SEQUENCE 974 AA; 107022 MW; 73B338F080FB451 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 974;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
| | | | |
Db 381 VLAALAA 387

RESULT 785
Q4QDN7 LEIMA PRELIMINARY; PRT; 974 AA.

ID Q4QDN7 LEIMA PRELIMINARY;
AC Q4QDN7;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Cation-transporting ATPase.
GN ORFNames=LmjF18.1520;
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.


```

OX NCBI_TaxID=5664;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=Friedlin;
RA Peacock C.S., Murphy L., Ivens A.C., Berriman M., Blackwell J.,
RA Smith D., Collins M., Foster N., Harris D., Oliver K., O'Neill S.,
RA Saunders D., Seeger K., Warren T., Rajandream M., and Barrell B.G.;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; CTO05257; CAJ04294.1; -; Genomic DNA.
DR InterPro; IPR006534; ATPase-IIIA_H.
DR InterPro; IPR001757; ATPase_EI-E2.
DR InterPro; IPR004014; Cation ATPase N.
DR InterPro; IPR005834; Dehal_like_hydro.
DR InterPro; IPR008250; EI-E2_ATPase_reg.
DR InterPro; IPR000695; H_ATPase.
DR Pfam; PF00690; Cation_ATPase_N.1.
DR Pfam; PF00122; EI-E2_ATPase.1.
DR Pfam; PF00702; Hydrolase.1.
DR PRINTS; PR00119; CATATPASE.
DR PRINTS; PR00120; HATPASE.
DR TIGRFAMs; TIGR01647; ATPase-IIIA_H.1.
DR TIGRFAMs; TIGR01494; ATPase_P-type.3.
DR PROSITE; PS00154; ATPASE_EI-E2; UNKNOWN_1.
KW ATP-binding; Hydrolase; Nucleotide-binding; Transmembrane.
SQ SEQUENCE 974 AA; 106963 MW; D3E58E408E0D3E57 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 974;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 VLAALAA 26
DB 381 VLAALAA 387

RESULT 786
Q8WZ29 NEUCR
ID Q8WZ29 NEUCR PRELIMINARY; PRT; 975 AA.
AC Q8WZ29
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein B24G3.020.
GN Name=B24G3.020;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysaselis M., Mauceli E., Biele C., Rudd S., Frisman D.,
RA Krystofova S., Rasmussen C., Metzner R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RN Nature 0:0-0(2003).
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.

```

```

RESULT 787
Q5TBP2 HUMAN
ID Q5TBP2 HUMAN PRELIMINARY; PRT; 984 AA.
AC Q5TBP2;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DE Retinoblastoma-associated factor 600 (RBAF600) (Fragment).
GN Name=RP5-1126H10.1; ORFNames=RP5-1126H10.1-006;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN NUCLEOTIDE SEQUENCE.
RA Sehra H.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL137127; CAL19274.1; -; Genomic DNA.
DR InterPro; IPR001920; Asp/Glu_rac.
DR PROSITE; PS00923; ASP_GLU_RACEMASE_1; UNKNOWN_1.
FT NON_TER 1 984
FT NON_TER 984
SQ SEQUENCE 984 AA; 106369 MW; E01C1F7ACA598448 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 984;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 VLAALAA 26
DB 940 VLAALAA 946

RESULT 788
Q7SHP3 NEUCR
ID Q7SHP3 NEUCR PRELIMINARY; PRT; 985 AA.
AC Q7SHP3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DE Hypothetical protein (Probable ribosomal elongation factor EF-2).
GN Name=NCU02572.1; Synonyms=B5K2.120;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Peizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysaselis M., Mauceli E., Biele C., Rudd S., Frisman D.,
RA Krystofova S., Rasmussen C., Metzner R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmani S.A.,
RA Desouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RN Nature 0:0-0(2003).
RP NUCLEOTIDE SEQUENCE.
RA Schulte U., Aign V., Hobeisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RP NUCLEOTIDE SEQUENCE.

```

```

RA German Neurospora genome project;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AABX01000002; EAA36467.1; -; Genomic DNA.
DR EMBL; BX842632; CAB76428.1; -; Genomic DNA.
DR GO; GO:0005525; F-GTP binding; IEA.
DR GO; GO:0003746; F-translation elongation factor activity; IEA.
DR GO; GO:0006412; F-protein biosynthesis; IEA.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR000795; ProtSyn_GTPbind.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PR00315; ELONGATNFCT.
DR TIGRFAMs; TIGR00231; small_GTP; 1.
KW Elongation factor; Hypothetical protein.
SQ SEQUENCE 985 AA; 110609 MW; DC4D6AFD0EC86117 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 985;
Best Local Similarity 100.0%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 14 WLLGGV 20
Db 562 WLLGGV 568

RESULT 789
Q4HBU4_9DEIO
ID Q4HBU4_9DEIO PRELIMINARY; PRT; 986 AA.
AC Q4HBU4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Dehydrogenase, E1 component (EC 1.2.4.2).
GN ORFNames=DgeODRAFT.1593;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=DSM 11300;
RC US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000002; EAL83487.1; -; Genomic DNA.
KW Oxidoreductase.
SQ SEQUENCE 986 AA; 108527 MW; 40291607DDC6B56EB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 986;
Best Local Similarity 100.0%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 14 WLLGGV 20
Db 562 WLLGGV 568

RESULT 789
Q4HBU4_9DEIO PRELIMINARY; PRT; 986 AA.
AC Q4HBU4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Dehydrogenase, E1 component (EC 1.2.4.2).
GN ORFNames=DgeODRAFT.1593;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=DSM 11300;
RC US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000002; EAA36467.1; -; Genomic DNA.
KW Oxidoreductase.
SQ SEQUENCE 986 AA; 108527 MW; 40291607DDC6B56EB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 986;
Best Local Similarity 100.0%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 14 WLLGGV 20
Db 562 WLLGGV 568

RESULT 789
Q4HBU4_9DEIO PRELIMINARY; PRT; 986 AA.
AC Q4HBU4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Dehydrogenase, E1 component (EC 1.2.4.2).
GN ORFNames=DgeODRAFT.1593;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
[1]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=DSM 11300;
RC US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
[2]
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300."
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000002; EAA36467.1; -; Genomic DNA.
KW Oxidoreductase.
SQ SEQUENCE 986 AA; 108527 MW; 40291607DDC6B56EB CRC64;

Query Match 5.9%; Score 7; DB 1; Length 1005;
Best Local Similarity 100.0%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

Qy 15 VLLGGV 21
Db 27 VLLGGV 33

RESULT 791
Q7M912_WOLSU

```

```

ID Q7M912_WOLSU PRELIMINARY; PRT; 1009 AA.
AC Q7M912;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RND PUMP PROTEIN.
DE Name=HEFC; OrderedLocusNames=WS1271;
OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX STRAIN=DSMZ 1740;
RX MEDLINE=22882897; PubMed=14500908; DOI=10.1073/pnas.132838100;
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
DR EMBL; BX571660; CAE10350.1; -; Genomic_DNA.
DR HSSP; P31224; 1IWG.
DR InterPro; IPR001036; Acrflvin_res.
DR Pfam; PF00873; Acr tran. 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR Complete proteome.
DR Complete proteome.
SQ SEQUENCE 1009 AA; 111567 MW; 49DEA7PFA04B41CB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1009;
Best Local Similarity 100.0%; Pred. No. 1.e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LGGVLA 22
|||||
Db 329 LGGVLA 335

RESULT 792
Q4R5S1_MACFA
ID Q4R5S1_MACFA PRELIMINARY; PRT; 1053 AA.
AC Q4R5S1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Testis cDNA, clone: QtsA-21352, similar to human retinoblastoma-
DE associated factor 600 (RBAF600).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheidae; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA International consortium for macaque cDNA sequencing, analysis;
RA "DNA sequences of macaque genes expressed in brain or testis and its
RA evolutionary implications.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Osada N., Hirata M., Tanuma R., Kusuda J., Hida M., Suzuki Y.,
RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
RT "Substitution rate and structural divergence of 5'UTR evolution:
RT Comparative analysis between human and cynomolgus monkey cDNAs.";
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB169472; BAE01554.1; -; mRNA.
DR SEQUENCE 1053 AA; 118565 MW; 9230B2703BC44D6B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1053;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
|||||

```

```

Db 532 VLAALAA 538

RESULT 793
Q7TUJ6_PROMM
ID Q7TUJ6_PROMM PRELIMINARY; PRT; 1099 AA.
AC Q7TUJ6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SNF2 related domain/DEAD/DEAH box helicase.
GN OrderedLocusNames=PMW2158;
OS Prochlorococcus marinus (strain MIT 9313).
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=74547;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22823698; PubMed=12917642; DOI=10.1038/nature01947;
RA Rocap G., Larimer F.W., Lamerdin J.E., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Heas W.R.,
RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation.";
RL Nature 424:1042-1047(2003).
DR EMBL; BX572101; CAE22332.1; -; Genomic_DNA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004386; F:helicase activity; IEA.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00176; SNF2_N; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
DR Complete proteome; Helicase.
DR Complete proteome; Helicase.
SQ SEQUENCE 1099 AA; 122557 MW; 36F13424772AF1E6 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1099;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGLL 90
|||||
Db 891 GKVLGLL 897

RESULT 794
Q54HL3_DICDI
ID Q54HL3_DICDI PRELIMINARY; PRT; 1200 AA.
AC Q54HL3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Argonaut-like protein.
GN Name=agnE; ORFNames=DD80220439;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Tungal B., Berriman M., Song J., Olsen R., Szafarski K., Xu Q.,
RA Mader M., Konfortov B.A., Rivero F.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nle X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,

```

RA Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsey R.,
 RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
 RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
 RA Louisedge H., Mungall K., Oliver K., Price C., Quail M.A.,
 RA Urushihara H., Hernandez J., Rabbins M., Steffen D., Sanders M.,
 RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
 RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
 RA Shauleky G., Schleicher M., Weinstein G., Rosenthal A., Cox E.C.,
 RA Chisholm R.L., Gibbs R., Loomis W.F., Plattner M., Kay R.R.,
 RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,
 RT "The genome of the social amoeba Dictyostelium discoideum";
 RL Nature 0:0-0(2005).
 CC -i- CAUTION: The sequence shown here is derived from an
 CC ENBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR ENBL; AAF101000306; EAL62770.1; -; Genomic DNA.
 SQ SEQUENCE 1200 AA; 136663 MW; 7DF17D809EADF934 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 1200;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 83 KGKVLGL 89
 |||||
 DB 83 KGKVLGL 89

 RESULT 795
 Q53785_STRAU PRELIMINARY; PRT; 1219 AA.
 ID Q53785; STRAU PRELIMINARY; PRT; 1219 AA.
 AC Q53785;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Transmembrane protein.
 GN Names=whiB2;
 OS Streptomyces aureofaciens.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1894;
 [1]
 NUCLEOTIDE SEQUENCE.
 RP MSBLINE=98234302; PubMed=9565673;
 RA Hometova D., Sprusansky O., Potuckova L., Sevcikova B., Novakova R.,
 RA Rezuchova B., Kormanec J.;
 RT "The gene downstream of streptomyces aureofaciens whiB encodes a large
 RT protein with proposed transmembrane localization, and is induced by
 RT glucose";
 RL Biochim. Biophys. Acta 1397:151-155(1998).
 DR ENBL; L22864; AAC1892.1; -; Genomic DNA.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF00535; Glycos_transf_2; 1.
 KW Transmembrane.
 SQ SEQUENCE 1219 AA; 128208 MW; 0CEA4C9617ACED32 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 1219;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 20 VLAALAA 26
 |||||
 DB 714 VLAALAA 720

 RESULT 796
 Q82DD9_STRAU PRELIMINARY; PRT; 1225 AA.
 ID Q82DD9; STRAU PRELIMINARY; PRT; 1225 AA.
 AC Q82DD9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative membrane protein.

GN OrderedLocusNames=SAV5043;
 OS Streptomyces avermitilis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 [1]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RC MEDLINE=22608306; PubMed=12692562; DOI=10.1038/nbt820;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 RA Sakaki Y., Hattori M., Omura S.;
 RT "Complete genome sequence and comparative analysis of the industrial
 RT microorganism Streptomyces avermitilis";
 RL Nat. Biotechnol. 21:526-531(2003).
 [2]
 NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RC MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism Streptomyces
 RT avermitilis: deducing the ability of producing secondary
 RT metabolites";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR ENBL; BA000030; BAC72755.1; -; Genomic DNA.
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF00535; Glycos_transf_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 1225 AA; 129125 MW; 2BFB00731D8BBD42 CRC64;

 Query Match 5.9%; Score 7; DB 2; Length 1225;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 20 VLAALAA 26
 |||||
 DB 724 VLAALAA 730

 RESULT 797
 Q6M4L4_CORGL PRELIMINARY; PRT; 1255 AA.
 ID Q6M4L4; CORGL PRELIMINARY; PRT; 1255 AA.
 AC Q6M4L4;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Superfamily I DNA or RNA helicase.
 GN OrderedLocusNames=cgl1985;
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 [1]
 NUCLEOTIDE SEQUENCE.
 RP STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RC PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;
 RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M.,
 RA Burkowski A., Dusch N., Eggeling L., Bickmanns B.J., Gaigalat L.,
 RA Goessmann A., Hartmann M., Huthmacher K., Kraemer R., Linke B.,
 RA McHardy A.C., Meyer F., Moeckel B., Pfeifferle W., Pühler A.,
 RA Rey D.A., Rueckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I.,
 RA Tauch A.;
 RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence
 RT and its impact on the production of L-aspartate-derived amino acids
 RT and vitamins";
 RL J. Biotechnol. 104:5-25(2003).
 DR ENBL; BX27153; CAF20145.1; -; Genomic DNA.
 DR GO; GO:0005524; F: ATP binding; IEA.
 DR GO; GO:0004003; F: ATP-dependent DNA helicase activity; IEA.
 DR GO; GO:0003677; F: DNA binding; IEA.
 DR GO; GO:0003723; F: RNA binding; IEA.
 DR GO; GO:0003724; F: RNA helicase activity; IEA.

DR GO; GO:000368; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR InterPro; IPR000212; UvrD-helicase.
 DR InterPro; IPR000606; Viral-helicase.
 DR Pfam; PF00580; UvrD-helicase; 2.
 DR Pfam; PF01443; Viral_helicase1; 1.
 KW Helicase.
 SQ SEQUENCE 1255 AA; 139540 MW; B1368E2F216439FB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 ADLEVT 11
 |||||
 Db 1184 ADLEVT 1190

RESULT 798

ID Q9KZL1_STRCO PRELIMINARY; PRT; 1268 AA.
 AC Q9KZL1;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE Putative integral membrane regulatory protein.
 GN OrderedLocusNames=SCO30333; ORFNames=SCO34.14c;
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
 RA Rabinowitz E., Rajandream M.A., Rutherford K.M., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 DR EMBL; AL939114; CAB88917.1; -; Genomic_DNA.
 DR InterPro; IPR001173; Glyco_trans_2.
 DR Pfam; PF00535; Glycos_transf_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 1268 AA; 133034 MW; 412B415829CE6CFD CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1268;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
 |||||
 Db 719 VLAALAA 725

RESULT 799

ID Q8C8X7_MOUSE PRELIMINARY; PRT; 1302 AA.
 AC Q8C8X7;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 DE clone:A930005B13 product:hypothetical ARM repeat structure containing
 DE protein, full insert sequence. (Fragment).

GN Name=1810009A16Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Mateu Y., Nikola I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX THE FANTOM Consortium;
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=C57BL/6J; TISSUE=Retina;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Iehii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama I., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito K., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK044277; BAC31850.1; -; mRNA.
DR Ensembl; ENSMUSG000006036; Mus musculus.
DR MGI; MGI:1916366; 1810009A16R1.
DR InterPro; IPR011989; ARM-like.
KW Hypothetical protein.
FT NON_TER 1302 1302
SQ SEQUENCE 1302 AA; 146813 MW; DD61F8FC5ECCB9F5 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 1302;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 20 VLAALAA 26
| | | | |
Db 170 VLAALAA 176
RESULT 800
Q4P748 USTMA PRELIMINARY; PRT; 1338 AA.
AC Q4P748;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=UM04065.1;
OS Ustilago maydis 521.
OC Eukaryota; Fungi; Basidiomycota; Ustilaginomycetes; Ustilago.
OC Ustilaginomycetidae; Ustilaginales; Ustilaginaceae; Ustilago.
OC NCBI_TaxID=237631;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=521;
RA Birren B., Nusbaum C., Abebe A., Abouelheil A., Adekoya E.,
RA Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachantang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom T., Blye J., Bogdavlavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
RA Jaffe D., Jones C., Kanai M., Kanat A., Kanvasselis M., Karlsson E.,
RA Kells C., Kieu A., Kisner P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokvitsang T., Lokvitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marbella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., McGhee T., Meldrim J., Meneus L.,
RA Mesirov J., Mihalev A., Mihova T., Mikelsen T., Mlenga V., Moru K.,
RA Moses J., Mulrain L., Munson G., Naylor J., Neves C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okawa O., O'leary S., Omotosho B.,
RA O'Neill K., Oseman S., Parker S., Perrin D., Phunkhang P., Piquani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,
RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,
RA Rutman M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,
RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,

RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,
RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,
RA Tenzing P., Tesfaye S., Theodore J., Thoulutang Y., Topham K.,
RA Towey S., Teamla T., Tsomo N., Vallee D., Vassiliev H.,
RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,
RA Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S.,
RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,
RA Zimmer A., Zody M., Lander E.;
RT "The genome sequence of Ustilago maydis.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -! SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -! SIMILARITY: Belongs to the cation transport ATPase (P-type)
CC family.
DR EMBL; AACFP01000141; EAK84990.1; -; Genomic_DNA.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR005834; Dehal_Like_hydro.
DR InterPro; IPR008250; E1-E2_ATPase_reg.
DR InterPro; IPR008544; P-ATPase-V.
DR Pfam; PF00122; E1-E2_ATPase; 1.
DR Pfam; PF00702; Hydrolyase; 1.
DR PRINTS; PR00119; CATATPASE.
DR TIGRFAMS; TIGR01494; ATPase_P-type; 2.
DR TIGRFAMS; TIGR01657; P-ATPase-V; 1.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN 1.
KW ATP-binding; Hydrolyase; Hypothetical protein; Nucleotide-binding;
KW Transmembrane.
SQ SEQUENCE 1338 AA; 147333 MW; E2F8D22BAC40EB5A CRC64;
Query Match 5.9%; Score 7; DB 2; Length 1338;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 17 LGGVLA 23
| | | | |
Db 504 LGGVLA 510
RESULT 801
Q06635 9ALPH PRELIMINARY; PRT; 1343 AA.
AC Q06635;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BICP4.
GN Name=BICP4;
OS Bovine herpesvirus 1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OC NCBI_TaxID=10320;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K22;
RX MEDLINE=92219360; PubMed=1313901;
RA Wirth U.V., Fraefel C., Vogt B., Vlcek C., Paces V., Schwyzer M.;
RT "Immediate-early RNA 2.9 and early RNA 2.6 of bovine herpesvirus 1 are
RT 3' coterminal and encode a putative zinc finger transactivator
RT protein".
RL J. Virol. 66:2763-2772 (1992).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=93172350; PubMed=8382298;
RA Fraefel C., Wirth U.V., Vogt B., Schwyzer M.;
RT "Immediate-early transcription over covalently joined genome ends of
RT bovine herpesvirus 1: the circ gene".
RL J. Virol. 67:1328-1333 (1993).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=94025583; PubMed=8212570;
RA Schwyzer M., Vlcek C., Menekse O., Fraefel C., Paces V.;

RT "Promoter, spliced leader, and coding sequence for BICP4, the largest
 RL of the immediate-early proteins of bovine herpesvirus 1.";
 RL Virology 197:349-357(1993).

DR EMBL; L14320; AAA03610.1; -; Unassigned DNA.
 DR GO; GO:0042025; C:host cell nucleus; IEA.
 DR GO; GO:0003700; P:transcription factor activity; IEA.
 DR GO; GO:0004563; P:transcriptional activator activity; IEA.
 DR GO; GO:0004594; P:positive regulation of transcription; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR005205; Herpes ICP4_C.
 DR InterPro; IPR005206; Herpes ICP4_N.
 DR Pfam; PF03585; Herpes ICP4_C; 1.
 DR Pfam; PF03584; Herpes ICP4_N; 1.
 SQ SEQUENCE 1343 AA; 136475 MW; 9A89EB1A8CFA37B5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1343;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYC 28
 |||||
 Db 851 AALAAAYC 857

RESULT 802

Q4LDG6_9ALPH
 ID Q4LDG6_9ALPH PRELIMINARY; PRT; 1343 AA.

AC Q4LDG6;
 DT 13-SEP-2005 (TRENBLrel. 31, Created)
 DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Immediate-early transactivator protein (Cell nucleus).

GN Name=BICP4;
 OS Bovine herpesvirus type 1.1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicelloviruses.
 OX NCBI_TaxID=79889;
 RN NUCLEOTIDE SEQUENCE.

RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=94025583; PubMed=8212570;
 RA Schwytzer M., Vlcek C., Menckse O., Fraefel C., Paces V.;
 RT "Promoter, spliced leader, and coding sequence for BICP4, the largest
 of the immediate-early proteins of bovine herpesvirus 1.";
 RL Virology 197:349-357(1993).

RL NUCLEOTIDE SEQUENCE.
 RP NUCLEOTIDE SEQUENCE.
 RA Schwytzer M., Paces V., Lettsworth G.J., Misra V., Buhk H.J.,
 RA Lowery D.E., Simard C., Bello L.J., Thiry E., Vlcek C.;
 RT "Complete DNA sequence of bovine herpesvirus 1.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ004801; CAA06139.1; -; Genomic DNA.
 DR EMBL; AJ004801; CAA06150.1; -; Genomic DNA.
 SQ SEQUENCE 1343 AA; 136475 MW; 9A89EB1A8CFA37B5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1343;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYC 28
 |||||
 Db 851 AALAAAYC 857

RESULT 803

Q5E034_VIBF1
 ID Q5E034_VIBF1 PRELIMINARY; PRT; 1354 AA.

AC Q5E034;
 DT 10-MAY-2005 (TRENBLrel. 30, Created)
 DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
 DE Superfamily I DNA helicases and helicase subunits.
 GN OrderedLocNames=VFA0542;
 OS Vibrio fischeri (strain ATCC 700601 / ES114).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.

RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15703294; DOI=10.1073/pnas.040990102;
 RA Ruby E.G., Urbanowski M., Campbell J., Dunn A., Faini M., Gunsalus R.,
 RA Lostroh P., Lupp C., McCann J., Millikan D., Schaefer A., Stabb E.,
 RA Stevens A., Visick K., Whistler C., Greenberg E.P.;
 RT "Complete genome sequence of *Vibrio fischeri*: a symbiotic bacterium
 with pathogenic congeners.";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:3004-3009(2005).
 DR EMBL; CP000021; AAW87612.1; -; Genomic DNA.
 DR GO; GO:0004386; F:helicase activity; IEA.
 DR InterPro; IPR006171; Toprim dom.
 DR InterPro; IPR006154; Toprim_sub.
 DR Pfam; PF01753; Toprim; 1.
 DR SMART; SM00493; TOPRIM; 1.
 KW Complete proteome; Helicase.
 SQ SEQUENCE 1354 AA; 150617 MW; D579B8A38C5B7B73 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1354;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGVVLGL 89
 |||||
 Db 1052 KGVVLGL 1058

RESULT 804

Q65565_9ALPH
 ID Q65565_9ALPH PRELIMINARY; PRT; 1385 AA.

AC Q65565;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Major capsid protein.
 OS Bovine herpesvirus 1.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.

OX NCBI_TaxID=10320;
 RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=Cooper;
 RX MEDLINE=95313343; PubMed=7793062;
 RA Vlcek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
 RA Lettsworth G.J., Schwytzer M.;
 RT "Nucleotide sequence analysis of a 30-kb region of the bovine
 herpesvirus 1 genome which exhibits a colinear gene arrangement with
 the UL21 to UL4 genes of herpes simplex virus.";
 RL Virology 210:100-108(1995).

DR EMBL; Z48053; CAA88114.1; -; Genomic DNA.
 DR PIR; S61236; S61236.
 DR HSSP; P06491; 1NO7.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; IPR000912; Herpes MCP.
 DR Pfam; PF03122; Herpes MCP; 1.
 DR PRINTS; PR00235; HSVCAFSIDMCP.
 SQ SEQUENCE 1385 AA; 149735 MW; 1B2117B7877060A3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1385;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLGGVLA 22
 |||||
 Db 288 LLGGVLA 294

RESULT 805

Q77CC1_9ALPH

```

ID Q77CC1_9ALPH PRELIMINARY; PRT; 1385 AA.
AC Q77CC1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Major capsid protein.
GN Name=UL19;
OS Bovine herpesvirus type 1.1.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=79889;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=95313343; PubMed=7793062;
RA Vilek C., Benes V., Lu Z., Kutish G.F., Paces V., Rock D.,
RA Letchworth G.J., Schwyer M.;
RT "Nucleotide sequence analysis of a 30-kb region of the bovine
RT herpesvirus 1 genome which exhibits a colinear gene arrangement with
RT the UL21 to UL4 genes of herpes simplex virus.";
RL Virology 210:100-108(1995).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Schwyer M., Paces V., Letchworth G.J., Misra V., Buhk H.J.,
RA Lowery D.E., Simard C., Bello L.J., Thiry E., Vilek C.;
RT "Complete DNA sequence of bovine herpesvirus 1.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ004801; CA006115.1; -; Genomic_DNA.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP; I.
DR PRINTS: PR00235; HSVCAPSIDMCP.
DR SEQUENCE 1385 AA; 149736 MW; 1B2117B7877060A3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1385;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 288 LLGGVLA 294

RESULT 806
Q6X230_9ALPH PRELIMINARY; PRT; 1391 AA.
AC Q6X230;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE UL19 major capsid protein.
GN ORFNames=BHV5-38;
OS Bovine herpesvirus 5.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=35244;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SV507/99;
RX MEDLINE=22850801; PubMed=12970418;
RX DOI=10.1128/JVI.77.19.10339-10347.2003;
RA Delhon G., Moraes M.P., Lu Z., Afonso C.L., Flores E.P., Weiblen R.,
RA Kutish G.F., Rock D.L.;
RT "Genome of bovine herpesvirus 5.";
RL J. Virol. 77:10339-10347(2003).
DR EMBL: AY261359; AAR86143.1; -; Genomic_DNA.
DR GO: GO:0019028; C: viral capsid; IEA.
DR GO: GO:0005198; F: structural molecule activity; IEA.
DR InterPro: IPR000912; Herpes_MCP.
DR Pfam: PF03122; Herpes_MCP; I.
DR PRINTS: PR00235; HSVCAPSIDMCP.
DR SEQUENCE 1391 AA; 149876 MW; 1FE03D553B80C846 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1392;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 LLGGVLA 22
Db 289 LLGGVLA 295

RESULT 807
Q9XER9_9ALPH PRELIMINARY; PRT; 1392 AA.
AC Q9XER9;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Transcription factor-like protein.
GN Name=HUA2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsae.
OX NCBI_TaxID=3702;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99214850; PubMed=10198637; DOI=10.1016/S1097-2765(00)80462-1;
RA Chen X., Meyerowitz E.M.;
RT "HUA1 and HUA2 are two members of the floral homeotic AGAMOUS
RT pathway.";
RL Mol. Cell 3:349-360(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98162728; PubMed=9501997;
RA Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT Sequence features of the regions of 1,191,918 bp covered by seventeen
RT physically assigned P1 clones.";
RL DNA Res. 4:401-414(1997).
DR EMBL: AF116556; AAD31171.1; -; mRNA.
DR EMBL: AB007648; BAB11170.1; -; Genomic DNA.
DR EMBL: AB006708; BAB11170.1; JOINED; Genomic_DNA.
DR PIR: T51947; T51947.
DR TRANSFAC: T05383;
DR InterPro: IPR000313; PWMF.
DR InterPro: IPR006569; RPR.
DR Pfam: PF00855; PWMF; 1.
DR SMART: SM00582; RPR; 1.
DR PROSITE: PS00812; PWMF; 1.
KW Nucleotide-binding.
SQ SEQUENCE 1392 AA; 151078 MW; D2AC1B0D2AA70F0F CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1392;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 44 GKPAIVP 50
Db 453 GKPAIVP 459

RESULT 808
Q6X1Z5_9ALPH PRELIMINARY; PRT; 1408 AA.
AC Q6X1Z5;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE BICP4 positive and negative gene regulator.
GN ORFNames=BHV5-61, BHV5-72;
OS Bovine herpesvirus 5.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.

```



```

OX NCBI_TaxID=35244;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=SV507/99;
RX MEDLINE=22850801; PubMed=12970418;
RA DOI=10.1128/JVI.77.19.10339-10347.2003;
RA Delhon G., Moraes M.P., Lu Z., Afonso C.L., Flores B.F., Weiblen R.,
RA Kutish G.F., Rock D.L.;
RT "Genome of bovine herpesvirus 5.";
RL J. Virol. 77:10339-10347(2003).
DR EMBL; AY261359; AAR86178.1; -; Genomic DNA.
DR EMBL; AY261359; AAR86167.1; -; Genomic DNA.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0016563; P:transcriptional activator activity; IEA.
DR GO; GO:0045941; P:positive regulation of transcription; IEA.
DR InterPro; IPR005205; Herpes ICP4_C.
DR InterPro; IPR005206; Herpes ICP4_N.
DR Pfam; PF03585; Herpes ICP4_C; 1.
DR Pfam; PF03584; Herpes ICP4_N; 1.
SQ SEQUENCE 1408 AA; 141784 MW; 90508869D9CED697 CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 2; Length 1408;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYC 28
Db 911 AALAAAYC 917

RESULT 809
IE18_PRVKA
ID IE18_PRVKA STANDARD; PRT; 1446 AA.
AC P33479;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Immediate-early protein IE180.
GN Name=IE;
OS Pseudorabies virus (strain Kaplan) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=33703;
RN [1]
RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
RX MEDLINE=91021039; PubMed=2171211;
RA Vleck C., Kozmik Z., Paces V., Schirm S., Schwytzer M.;
RT "Pseudorabies virus immediate-early gene overlaps with an oppositely
oriented open reading frame: characterization of their promoter and
enhancer regions.";
RL Virology 179:365-377(1990).
CC -!- FUNCTION: This IE protein is a multifunctional protein capable of
migrating to the nucleus, binding to DNA, trans-activating other
viral genes, and autoregulating its own synthesis.
CC -!- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -!- PTM: A long stretch of serine residues may be a major site of
phosphorylation.
CC -!- SIMILARITY: Belongs to the herpesviruses ICP4/IE140/IE180 family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
CC EMBL; M34651; AAA47470.1; -; Genomic DNA.
CC PIR; A45344; A45344.
CC InterPro; IPR005205; Herpes ICP4_C.
CC InterPro; IPR005206; Herpes ICP4_N.
CC Pfam; PF03585; Herpes ICP4_C; 1.
CC Pfam; PF03584; Herpes ICP4_N; 1.
CC DNA-binding; Early protein; Nuclear protein; Phosphorylation;
Transcription; Transcription regulation.
KW

FT COMBIAS 347 354 Poly-Ser.
FT COMBIAS 379 397 Poly-Ser.
SQ SEQUENCE 1446 AA; 148642 MW; 81F43A3DE3DDA068 CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 1; Length 1446;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYC 28
Db 991 AALAAAYC 997

RESULT 810
QSP75_9ALPH
ID QSP75_9ALPH PRELIMINARY; PRT; 1446 AA.
AC QSP75;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Immediate-early protein IE180.
GN Name=IE180(TRS); Synonyms=IE180(IRS);
OS Suid herpesvirus 1 (Pseudorabies virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14671123; DOI=10.1128/JVI.78.1.424-440.2004;
RA Klupp B.G., Hengartner C.J., Mettenleiter T.C., Enquist L.W.;
RT "Complete, annotated sequence of the pseudorabies virus genome.";
RL J. Virol. 78:424-440(2004).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK001744; DAA02213.1; -; Genomic DNA.
DR EMBL; BK001744; DAA02202.1; -; Genomic DNA.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0016563; P:transcriptional activator activity; IEA.
DR GO; GO:0045941; P:positive regulation of transcription; IEA.
DR InterPro; IPR005205; Herpes ICP4_C.
DR InterPro; IPR005206; Herpes ICP4_N.
DR Pfam; PF03585; Herpes ICP4_C; 1.
DR Pfam; PF03584; Herpes ICP4_N; 1.
SQ SEQUENCE 1446 AA; 148639 MW; 81F43A3DE3DDA068 CRC64;

Query Match
Best Local Similarity 5.9%; Score 7; DB 2; Length 1446;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AALAAAYC 28
Db 991 AALAAAYC 997

RESULT 811
Q8JL63_9ALPH
ID Q8JL63_9ALPH PRELIMINARY; PRT; 1454 AA.
AC Q8JL63;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Immediate-early protein IE180.
OS Suid herpesvirus 1 (Pseudorabies virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TNL;
RX MEDLINE=22763901; PubMed=12881635; DOI=10.1023/A:1020959521745;
RA Ou C.J., Huang C., Wong M.L., Chang T.J.;
RT "Suppression of Promoter Activity of the LAT Gene by IE180 of
Pseudorabies Virus.";

```

RL Virus Genes 25:227-239(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TNL;
RX MEDLINE=22763902; PubMed=12881636; DOI=10.1023/A:1020915706724;
RA Ou C.-J., Wang M.L., Chang T.J.;
RT "A TEF-1-Element is Required for Activation of the Promoter of Pseudorabies Virus Glycoprotein X Gene by IE180.";
RL Virus Genes 25:241-253(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=TNL;
RA Chang T.-J., Ou C.-J.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF525264; AAM98908.1; -; Genomic DNA.
DR GO; GO:0042025; C:host cell nucleus; IEA.
DR GO; GO:0016563; F:transcriptional activator activity; IEA.
DR GO; GO:0045941; P:positive regulation of transcription; IEA.
DR InterPro; IPR005205; Herpes_ICP4_C.
DR Pfam; PF03585; Herpes_ICP4_C; 1.
DR Pfam; PF03584; Herpes_ICP4_N; 1.
SQ SEQUENCE 1454 AA; 150008 MW; B4B51501DABA1A94 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAYC 28
Db 998 AALAAAYC 1004
|||||
NUCLEOTIDE SEQUENCE.
PRT; 1461 AA.

RESULT 812
IE18 PRIVIF STANDARD; PRT; 1461 AA.
AC P11675;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Immediate-early protein IE180.
GN Name=IE;
OS Pseudorabies virus (strain Indiana-Funkhauser / Becker) (PRV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31523;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=89315207; PubMed=2546124;
RA Cheung A.K.;
RT "DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies virus.";
RL Nucleic Acids Res. 17:4637-4646(1989).
RN [2]
RP SEQUENCE REVISION.
RA Cheung A.K.;
RL Submitted (NOV-1989) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This IE protein is a multifunctional protein capable of migrating to the nucleus, binding to DNA, trans-activating other
CC -1- SUBCELLULAR LOCATION: Nucleus of infected cells.
CC -1- PTM: A long stretch of serine residues may be a major site of phosphorylation.
CC -1- SIMILARITY: Belongs to the herpesviruses ICP4/IE140/IE180 family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
DR EMBL; X15120; CAA33214.1; -; Genomic DNA.
DR PIR; S04713; EDSEIF.

DR InterPro; IPR005205; Herpes_ICP4_C.
DR InterPro; IPR005206; Herpes_ICP4_N.
DR Pfam; PF03585; Herpes_ICP4_C; 1.
DR Pfam; PF03584; Herpes_ICP4_N; 1.
KW DNA-binding; Early protein; Nuclear protein; Phosphorylation;
FT COMPBIAS 390 405 Poly-Ser.
FT COMPBIAS 958 966 Poly-Ser.
SQ SEQUENCE 1461 AA; 149834 MW; 7F31E7ABE403B208 CRC64;

Query Match 5.9%; Score 7; DB 1; Length 1461;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAAAYC 28
Db 1004 AALAAAYC 1010
|||||
NUCLEOTIDE SEQUENCE.
PRT; 1634 AA.

RESULT 813
Q511KO MAGGR PRELIMINARY; PRT; 1634 AA.
ID Q511KO MAGGR PRELIMINARY;
AC Q511KO;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MGI0526.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthae.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nuebaum C., Abebe A., Abouelleil A., Adekoya E., Alt-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S., Arachchi H., Armbruster J., Bachanteang P., Baldwin J., Barry A., Bayul T., Blitshsteyn B., Bloom T., Blye J., Boguslavsky L., Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N., Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M., Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C., David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P., Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R., Erickson J., Farina A., Faro S., Ferreira P., Fischer H., Fitzgerald M., Foley K., Gage D., Galeagan J., Gearin G., Gnerre S., Gnikre A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N., Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H., Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I., Jaffe D., Jones C., Kamal M., Kamat A., Kamyseselis M., Karlsson E., Kells C., Kleu A., Kisner P., Kodira C., Kulbokas E., Labutti K., Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T., Lindblad-toh K., Liu X., Lokvitsang T., Lokvitsang Y., Lucien O., Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J., Manning J., Marabella R., Maru K., Matthews C., Mauceli E., McCarthy M., McDonough S., Mcghee T., Meldrim J., Meneus L., Meisovich J., Mihalev A., Mihova T., Mikkelson T., Mlenga V., Moru K., Moses J., Mulrain L., Munson G., Naylor J., Neues C., Nguyen C., Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C., Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B., O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Piquani B., Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C., Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P., Rutman M., Schupbach R., Seaman C., Settipalli S., Sharpe T., Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C., Spencer B., Stalker J., Stange-thomann N., Stavropoulos S., RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P., Tenzing P., Tefaye S., Theodore J., Thoultsang Y., Topham K., RA Towey S., Teamlia T., Tsomo N., Vallee D., Vassiliev H., Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T., Wangdi T., Whittaker C., Wilkinson J., Wu Y., Wyman D., Yadav S., RA Yang S., Yang X., Yeager S., Yee B., Young G., Zainoun J., Zembeck L., Zimner A., Zody M., Lander E.;

RT "The genome sequence of Magnaporthe grisea.";
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [5]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [6]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [7]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [8]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [9]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [10]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [11]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [12]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [13]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [14]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [15]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [16]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [17]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [18]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [19]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [20]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [21]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [22]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [23]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [24]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [25]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
 RN [26]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Zhu H., Blackmon B.;

SQ SEQUENCE 1764 AA; 194773 MW; 5AFF85A32114B227 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1764;

Best Local Similarity 100.0%; Pred. No. 1.6e+03;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 VIGLLQR 92

DB 624 VIGLLQR 630

RESULT 815

Q6C6Q4_YARLI PRELIMINARY; PRT; 1820 AA.

AC Q6C6Q4;

DT 25-OCT-2004 (TRENBLrel. 28, Created)

DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)

DE Similar to sp|P32528 Saccharomyces cerevisiae YBR208c DURI_2 urea

DE amidolyase.

GN OrderedLocustNames=YALIOE07271g;

OS Yarrowia lipolytica (Candida lipolytica).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Dipodascaceae; Yarrowia.

OX NCBI_TaxID=4952;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

EX PubMed=15229592; DOI=10.1038/nature02579;

RA Dujon B., Sherman D., Montigny J., Marck C., Neugebauer S., Tallia E.,

RA Lafontaine I., de Montigny J., Marck C., Neugebauer S., Tallia E.,

RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,

RA Barnay S., Blanchin S., Beckerich J.-M., Beyne E., Bleykasten C.,

RA Bolesame A., Boyer J., Cattolico L., Confanioli P., de Daruvar A.,

RA Despons L., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,

RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,

RA Nicaud J.-M., Nikolaki M., Oztas S., Ozier-Kalogeropoulos O.,

RA Pellenz S., Potier S., Richard G.-F., Straub M.-L., Suleau A.,

RA Swennen D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,

RA Zeniou-Meyer M., Zivanovic Y., Bolotin-Fukuhara M., Thierry A.,

RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,

RA Wincker P., Souciet J.-L.;

RT "Genome evolution in yeasts.";

RL Nature 430:35-44(2004).

DR EMBL; CR382131; CAG79240.1; -; Genomic DNA.

DR GO; GO:0004040; F:amidase activity; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR GO; GO:0009374; F:biotin binding; IEA.

DR GO; GO:0016874; F:ligase activity; IEA.

DR GO; GO:0016829; F:lyase activity; IEA.

DR GO; GO:0008152; P:metabolism; IEA.

DR InterPro; IPR003833; Allophan hydrol.

DR InterPro; IPR003778; Allophan hydrol.

DR InterPro; IPR000120; Amidase.

DR InterPro; IPR011761; ATP GRASP.

DR InterPro; IPR001882; Biotin BS.

DR InterPro; IPR005482; Biotin carb C.

DR InterPro; IPR000089; Biotin lipoyl.

DR InterPro; IPR005481; CPase L.N.

DR Pfam; PF02682; AHS1; 1.

DR Pfam; PF02626; AHS2; 1.

DR Pfam; PF01425; Amidase; 1.

DR Pfam; PF02785; Biotin carb C; 1.

DR Pfam; PF00364; Biotin lipoyl; 1.

DR Pfam; PF00289; CPase L chain; 1.

DR Pfam; PF02786; CPase_L_D2; 1.

DR TIGRPFAMs; TIGR00724; urea_amine_rel; 1.

DR PROSITE; PS00975; ATP GRASP; 1.

DR PROSITE; PS00979; BC; 1.

DR PROSITE; PS00188; BIOTIN; 1.

DR PROSITE; PS00968; BIOTINYL_LIPOYL; 1.

DR PROSITE; PS00866; CPSASE 1; UNKNOWN 1.
 DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
 KW Complete proteome; Lyase; 0; Indels 0; Gaps 0;
 SQ SEQUENCE 1820 AA; 199755 MW; B9F3329D0560786B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 1820;
 Best Local Similarity 100.0%; Pred. No. 1.7e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 98 AVIEPPIV 104
 |||||
 Db 1814 AVIEPPIV 1820

RESULT 816
 Q8CHF3_MOUSE
 ID Q8CHF3_MOUSE PRELIMINARY; PRT; 2056 AA.
 AC Q8CHF3;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE KIAA0462 protein (Fragment).
 GN Name=181009A16Rik; Synonyms=mKIAA0462;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Hara Y., Nagase T.,
 RA Chara O., Koga H.;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB093242; BAC41426.2; -; mRNA.
 DR Ensembl; ENSMUSG0000066036; Mus musculus.
 DR MGI; MGI:1916366; 181009A16Rik.
 DR GO; GO:0004842; P-ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P-ubiquitin cycle; IEA.
 DR InterPro; IPR001920; Asp/Glu_rac.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR003126; Znf_Nrecognin.
 DR Pfam; PF02207; zf-UBR1; 1.
 DR SMART; SM00396; Znf UBR1; 1.
 DR PROSITE; PS00923; ASP_GLU_RACEMASE 1; UNKNOWN 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 FT NON TER 1
 FT NON TER 2056 2056
 SQ SEQUENCE 2056 AA; 224200 MW; DA8F113B4C3A76C3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 2056;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 2016 VLAALAA 2022

RESULT 817
 Q5T4S9_HUMAN
 ID Q5T4S9_HUMAN PRELIMINARY; PRT; 2058 AA.
 AC Q5T4S9;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Retinoblastoma-associated factor 600 (RBAF600) (Fragment).
 GN Name=RP5-1126H10.1; ORFNames=RP5-1126H10.1-007;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
 RA Nakajima D., Nomura N., Ohara O.;
 RA "Characterization of cDNA clones in size-fractionated cDNA libraries
 from human brain.";
 RL DNA Res. 4:345-349(1997).
 DR EMBL; AB007931; BAA32307.1; -; mRNA.
 DR PIR; T00076; T00076.
 FT NON TER 1
 FT NON TER 2058 2058
 SQ SEQUENCE 2058 AA; 225005 MW; DB8F0FCAAA2C2FCB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 2058;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 425 VLAALAA 431

RESULT 819
 Q4FWP4_LEIMA
 ID Q4FWP4_LEIMA PRELIMINARY; PRT; 2954 AA.

RP NUCLEOTIDE SEQUENCE.
 RA Brown A.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sehra H.;
 RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL357564; CA120972.1; -; Genomic DNA.
 DR EMBL; AL37127; CA119268.1; -; Genomic DNA.
 DR EMBL; AL37127; CA120972.1; JOINED; Genomic DNA.
 DR EMBL; AL357564; CA119268.1; JOINED; Genomic DNA.
 DR GO; GO:0004842; P-ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0006512; P-ubiquitin cycle; IEA.
 DR InterPro; IPR001920; Asp/Glu_rac.
 DR InterPro; IPR007087; Znf_C2H2.
 DR InterPro; IPR003126; Znf_Nrecognin.
 DR Pfam; PF02207; zf-UBR1; 1.
 DR SMART; SM00396; Znf UBR1; 1.
 DR PROSITE; PS00923; ASP_GLU_RACEMASE 1; UNKNOWN 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 FT NON TER 1
 FT NON TER 2058 2058
 SQ SEQUENCE 2058 AA; 225005 MW; DB8F0FCAAA2C2FCB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 2058;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 2018 VLAALAA 2024

RESULT 818
 O75050_HUMAN
 ID O75050_HUMAN PRELIMINARY; PRT; 2276 AA.
 AC O75050; O9UG82;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE KIAA0462 protein (Fragment).
 GN Name=KIAA0462;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=98116662; PubMed=9455484;
 RA Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
 RA Nakajima D., Nomura N., Ohara O.;
 RA "Characterization of cDNA clones in size-fractionated cDNA libraries
 from human brain.";
 RL DNA Res. 4:345-349(1997).
 DR EMBL; AB007931; BAA32307.1; -; mRNA.
 DR PIR; T00076; T00076.
 FT NON TER 1
 FT NON TER 2276 2276
 SQ SEQUENCE 2276 AA; 255392 MW; D4726D76PFA257F0 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 2276;
 Best Local Similarity 100.0%; Pred. No. 2e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALAA 26
 |||||
 Db 425 VLAALAA 431

RESULT 819
 Q4FWP4_LEIMA
 ID Q4FWP4_LEIMA PRELIMINARY; PRT; 2954 AA.

AC Q4FWP4;
 DT 13-SEP-2005 (TRENBLrel. 31, Created)
 DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=LMJ_1155;
 OS Leishmania major.
 OC Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Friedlin;
 EX PubMed=16020728; DOI=10.1126/science.1112680;
 RA Ivens A.C., Peacock C.S., Worthey E.A., Murphy L., Aggarwal G.,
 RA Bertram M., Sisk E., Rajandream M.A., Adlem E., Aert R., Anupama A.,
 RA Apostolou Z., Attipoe P., Bason N., Bauser C., Beck A., Beverley S.M.,
 RA Bianchetti G., Borzym K., Bothe G., Bruschi C.V., Collins M.,
 RA Cadag E., Ciarloni L., Clayton C., Coulson R.M., Cronin A., Cruz A.K.,
 RA Davies R.M., De Gaudenzi J., Dobson D.E., Duisterhoef A.,
 RA Fazaelina G., Fosker N., Frasch A.C., Fraser A., Fuchs M., Gabel C.,
 RA Goble A., Goffeau A., Harris D., Hertz-Fowler C., Hilbert H., Horn D.,
 RA Huang Y., Klages S., Knights A., Kube M., Larke N., Litvin L.,
 RA Lord A., Louie T., Marra M., Masuy D., Matthews K., Michaeli S.,
 RA Mottram J.C., Muller-Auer S., Munden H., Nelson S., Norbertzak H.,
 RA Oliver K., O'Neill S., Pentony M., Pohl T.M., Price C., Purnelle B.,
 RA Quail M.A., Rabinowitsch E., Reinhardt R., Rieger M., Rinta J.,
 RA Robben J., Robertson L., Ruiz J.C., Rutter S., Saunders D.,
 RA Schafer M., Schein J., Schwartz D.C., Seeger K., Seyler A., Sharp S.,
 RA Shin H., Sivam D., Squares R., Squares S., Tosato V., Vogt C.,
 RA Volckaert G., Wambutt R., Warren T., Wedler H., Woodward J., Zhou S.,
 RA Zimmermann W., Smith D.F., Blackwell J.M., Stuart K.D., Barrell B.,
 RA Myler P.J.;
 RL "The Genome of the Kinetoplastid Parasite, Leishmania major.";
 RL Science 309:436-442(2005).
 DR EMBL: CP000081; AAZ14449.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 2954 AA; 311344 MW; 60FF96F29E3D34944 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 2954;
 Best Local Similarity 100.0%; Pred. No. 2.4e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GVLAALA 25
 |||||
 Db 2714 GVLAALA 2720

RESULT 820
 POLG HCVJ6
 ID POLG HCVJ6 STANDARD; PRT; 3032 AA.
 AC P26660;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Genome polyprotein [Contains: Core protein p19 (Capsid protein C)
 DE (p21); Envelope glycoprotein E1 (gp32) (gp35); Envelope glycoprotein
 DE E2 (NS1) (gp68) (gp70); p7; Protease NS2-3 (EC 3.4.22.-) (p23); Serine
 DE protease/NTase/helicase NS3 (EC 3.4.21.98) (3.6.1.15) (EC 3.6.1.-)
 DE (Hepacivirin) (NS3P) (p70); Nonstructural protein 4A (NS4A) (p8);
 DE Nonstructural protein 4B (NS4B) (p27); Nonstructural protein 5A (NS5A)
 DE (p56); RNA-directed RNA polymerase (EC 2.7.7.48) (NS5B) (p68)].
 OS Hepatitis C virus (isolate HC-J6) (HCV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OC NCBI_TaxID=11113;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC RNA].
 RX MEDLINE=9204440; PubMed=1658196;
 RA Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H., Machida A.,
 RA Miyakawa Y., Mayumi M.;
 RT "Nucleotide sequence of the genomic RNA of hepatitis C virus isolated
 RT from a human carrier: comparison with reported isolates for conserved
 RT and divergent regions.";

RL J. Gen. Virol. 72:2697-2704(1991).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 2442-3011.
 RX PubMed=15746101; DOI=10.1074/jbc.M413410200;
 RA Biswal B.K., Cherney M.M., Wang M., Chan L., Yannopoulos C.G.,
 RA Bilimoria D., Nicolas O., Bedard J., James M.N.;
 RT "Crystal structures of the RNA-dependent RNA polymerase genotype 2a of
 RT hepatitis C virus reveal two conformations and suggest mechanisms of
 RT inhibition by non-nucleoside inhibitors.";
 RL J. Biol. Chem. 280:18202-18210(2005).
 CC -!- FUNCTION: Core protein packages viral RNA to form a viral
 CC nucleocapsid, and promotes virion budding. Modulates viral
 CC translation initiation by interacting with HCV IRES and 40S
 CC ribosomal subunit. Also regulates many host cellular functions
 CC such as signaling pathways and apoptosis. Prevents the
 CC establishment of cellular antiviral state by blocking the
 CC interferon-alpha/beta (IFN-alpha/beta) and IFN-gamma signaling
 CC pathways and by inducing human STAT1 degradation. Plays an
 CC important role in virus-mediated cell transformation leading to
 CC hepatocellular carcinomas. Interacts with, and activates STAT3
 CC leading to cellular transformation. May repress the promoter of
 CC p53, and sequester CREB3 and SP10 isoform 3/Sp10b in the
 CC cytoplasm. Also represses cell cycle negative regulating factor
 CC CKN1A, thereby interrupting an important check point of normal
 CC cell cycle regulation. Targets transcription factors involved in
 CC the regulation of inflammatory responses and in the immune
 CC response: suppresses NK-kappaB activation, and activates AP-1.
 CC Mediates apoptotic pathways through association with TNF-type
 CC receptors TNFRSF1A and TNFR, although its effect on death
 CC receptors-induced apoptosis remains controversial. Enhances TRAIL
 CC mediated apoptosis, suggesting that it might play a role in
 CC immune-mediated liver cell injury. Secreted core protein is able
 CC to bind C1QRI at the T-cell surface, resulting in down-regulation
 CC of T-lymphocytes proliferation. May transactivate human MYC. Rous
 CC sarcoma virus LTR, and SV40 promoters. May suppress the human FOS
 CC and HIV-1 LTR activity. May alter lipid metabolism by interacting
 CC with hepatocellular proteins involved in lipid accumulation and
 CC storage (By similarity).
 CC -!- FUNCTION: Envelope glycoproteins E1 and E2 are involved in virus
 CC attachment to the host cell as well as in virus endocytosis and
 CC fusion with host membrane. E2 inhibits human PKR activation,
 CC preventing the establishment of an antiviral state (By
 CC similarity).
 CC -!- FUNCTION: p7 seems to be a hexameric ion channel protein
 CC (viroporin) and is inhibited by the antiviral drug amantadine.
 CC Also inhibited by long-alkyl-chain iminosugar derivatives.
 CC Essential for infectivity (By similarity).
 CC -!- FUNCTION: Protease NS2-3, which is a putative cysteine protease,
 CC is responsible for the autocatalytic cleavage of NS2-NS3. Seems to
 CC undergo self-inactivation following maturation (By similarity).
 CC -!- FUNCTION: NS3 displays three distinct enzymatic activities: serine
 CC protease, NTPase and RNA helicase. NS3 serine protease, in
 CC association with NS4A, is responsible for the cleavages of NS3-
 CC NS4A, NS4A-NS4B, NS4B-NS5A and NS5A-NS5B. NS3 RNA helicase binds
 CC to RNA and unwinds dsRNA in the 3' to 5' direction (By
 CC similarity). NS3/NS4A complex also prevents phosphorylation of
 CC human IRF3, thus preventing the establishment of dsRNA induced
 CC antiviral state.
 CC -!- FUNCTION: NS4B may induce a specific membrane alteration that
 CC serves as a scaffold for the virus replication complex. This
 CC membrane alteration gives rise to the so called ER-derived
 CC membranous web that contains the replication complex (By
 CC similarity).
 CC -!- FUNCTION: NS5A is a component of the replication complex.
 CC Downregulates viral IRES translation initiation. Mediates
 CC interferon resistance, presumably by interacting with and
 CC inhibiting human PKR/p38. The hyperphosphorylated form of NS5A is
 CC an inhibitor of viral replication (By similarity).
 CC -!- FUNCTION: NS5B is a RNA-dependent RNA polymerase that plays an
 CC essential role in the virus replication (By similarity).
 CC -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
 CC precursor polyprotein, commonly with Asp or Glu in the P6
 CC position, Cys or Thr in P1 and Ser or Ala in P1'.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).
-1- COFACTOR: Binds 1 zinc ion per NS3 protease domain. Binds 1 zinc ion per NS5A N-terminal domain (By similarity).
-1- ENZYME REGULATION: Activity of protease NS2-3 is dependent on zinc ions and completely inhibited by EDTA. Serine protease NS3 is also activated by zinc ions (By similarity).
-1- SUBUNIT: Core protein is a homodimer that binds the C-terminal part of E1. Interacts with numerous cellular proteins. Interacts with human STAT1, inducing its degradation and human STAT3, constitutively activating it. Associates with human LTR and TNFRSF1A receptors and possibly induces apoptosis. Binds to human SP110 isoform 3/Sp110b, HNRPK, CLOR1, YWAEF, DDX3X, APOA2 and RXRA proteins. Interacts with human CREB3 nuclear transcription protein, triggering cell transformation. May interact with human p53. Also binds human cytochromes KRT8, KRT18, KRT19 and VIM (vimentin). E1 and E2 glycoproteins form a heterodimer that binds to human LDLR, CD81 and SCARB1 receptors, but this binding is not sufficient for infection, some additional liver specific cofactors may be needed. E2 binds and inhibits human PRKR. p7 forms a homohexamer. NS2 forms a homodimer and interacts with all other nonstructural (NS) proteins. NS4A interacts with NS3 serine protease and stabilizes it. NS3-NS4A complex is essential for the activation of the latter and allows membrane anchorage of NS3. NS4B forms a homodimer and interacts with all other NS proteins. NS5A forms a homodimer and interacts with all other nonstructural (NS) proteins. NS5A also interacts with human PRKR, GRB3, PIK3R1 and with most Src-family kinases. NS5B is a homooligomer (By similarity).
-1- SUBCELLULAR LOCATION: The virion assembly and budding occurs from the ER membrane. The N-terminal transmembrane domains of the glycoprotein probably possess an endoplasmic reticulum (ER) signal sequence function in their C-termini, leading the nascent polypeptide to the ER membrane. After cleavage by host signal peptidase, these ER signal peptides retain at the C-terminus of the concerned proteins (core, E1, E2 and p7), serving as ER membrane anchors. Core protein is cytoplasmic. It is also located on mitochondrial and endoplasmic reticulum membranes and at the surface of lipid droplets. A minor proportion is present in the nucleus. An unknown proportion is secreted. E1, E2, NS2 and NS4B are integral ER membrane proteins. The C-terminal transmembrane domains of envelope glycoproteins E1 and E2 form a hairpin structure before cleavage by host signal peptidase. A reorientation of the second hydrophobic stretch occurs after cleavage producing a single reoriented transmembrane domain. These events explain the final topology of these proteins. ER retention of E1 and E2 is leaky and, in overexpression conditions, a small fraction of both proteins reaches the plasma membrane. NS3 is associated to the ER membrane through its binding to NS4A. Membrane insertion of the membrane-anchored proteins NS4A, NS5A and NS5B occurs after processing by the NS3 protease. NS5A is perinuclear. A fraction of p7 localizes to the plasma membrane (By similarity).
-1- DOMAIN: The transmembrane regions of envelope E1 and E2 glycoproteins are involved in heterodimer formation, ER localization, and assembly of these proteins. Envelope E2 glycoprotein contain two highly variable regions called hypervariable region 1 and 2 (HVR1 and HVR2) and two CD81-binding sites. HVR1 is implicated in the SCARB1-mediated cell entry. HVR2 and CD81-binding sites may be involved in sensitivity and/or resistance to IFN-alpha therapy (By similarity).
-1- DOMAIN: The N-terminal one-third of serine protease NS3 contains the protease activity. This region contains a zinc atom that does not belong to the active site, but may play a structural rather than a catalytic role. This region is essential for the activity of protease NS2-3, maybe by contributing to the folding of the latter. The helicase activity is located in the C-terminus (By similarity).
-1- PTM: Specific enzymatic cleavages in vivo yield mature proteins. The structural proteins, core, E1, E2 and p7 are produced by proteolytic processing by host signal peptidases. The core protein is synthesized as a 21 kDa precursor which is retained in the ER membrane through a hydrophobic propeptide, which is cleaved to

release the 19 kDa mature core protein. The other proteins are cleaved by the viral proteases (By similarity).
-1- PTM: Envelope E1 and E2 glycoproteins are highly N-glycosylated (By similarity).
-1- PTM: Core protein is phosphorylated by host PKC and PKA (By similarity).
-1- PTM: NS5A is phosphorylated in a basal form termed p56. p58 is an hyperphosphorylated form of p56. p56 and p58 coexist in the cell in roughly equivalent amounts. Hyperphosphorylation is dependent on the presence of NS4A in cis. Human AKT1, RPS6KB1/p70S6K, MAP2K1/MEK1 and MAP2K6/MKK6 kinases may be responsible for NS5A phosphorylation (By similarity).
-1- MISCELLANEOUS: Cell culture adaptation of the virus leads to mutations in NS5A, reducing its inhibitory effect on replication (By similarity).
-1- MISCELLANEOUS: The virion is a nucleocapsid covered by a lipoprotein envelope. The nucleocapsid is composed of the core protein forming an internal icosahedral coat that encapsidates the genomic RNA. The envelope contains two proteins: the envelope glycoproteins E1 and E2.
-1- MISCELLANEOUS: Core protein exerts viral interference on hepatitis B virus when HCV and HBV coinfect the same cell, by suppressing HBV gene expression, RNA encapsidation and budding (By similarity).
-1- SIMILARITY: Contains 1 peptidase C18 domain.
-1- SIMILARITY: Contains 1 peptidase S29 domain.
-1- CAUTION: There is a doubt concerning the orientation of the N-terminus of NS4B, which could be luminal with a 5th transmembrane segment. The C-terminus of NS2 may be luminal with a fourth transmembrane segment.

Query Match 5.9%; Score 7; DB 1; Length 3032;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1699 PDKEVLY 1705
|||||

RESULT 821
Q991B4_9HEPC
ID Q991B4_9HEPC PRELIMINARY; PRT; 3032 AA.
AC Q991B4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE
MEDLINE=21316767; PubMed=1424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Wakita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient.";
RL J. Med. Virol. 64:334-339(2001).
DR EMBL; AB047643; BAB32876.1; -; Genomic_RNA.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q991B4; 1032-1660, 2442-3007.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.

```
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3032 AA; 328955 MW; 04D4CD3B5B40AFCE CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3032;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1699 PDKEVLY 1705

RESULT 822
Q909A9_9HEPC
ID Q909A9_9HEPC PRELIMINARY; PRT; 3032 AA.
AC Q909A9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21361470; PubMed=11468731; DOI=10.1002/jmv.1073;
RA Kurihara K., Ishiyama N., Nishiyama Y., Fukushi S., Kageyama T.,
RA Katayama K., Miura S.;
RT "Molecular characterization of hepatitis C virus genotype 2a from the
RT entire sequences of four isolates.";
RL J. Med. Virol. 64:466-475(2001).
DR EMBL; AF169003; AAF25611.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q909A9; 1032-1660, 2442-3002.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003969; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
```

```
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3032 AA; 329397 MW; BAF67F4730C124B9 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3032;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1699 PDKEVLY 1705

RESULT 823
Q99IB3_9HEPC
ID Q99IB3_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q99IB3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Wakita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient.";
RL J. Med. Virol. 64:334-339(2001).
DR EMBL; AB047644; BAB32877.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CWK.
DR SMR; Q99IB3; 1033-1661, 2444-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; F:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
```

```
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol DS ps.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptidein.
SQ SEQUENCE 3033 AA; 328908 MW; B306821483E70CB3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 824
Q99IB6_9HEPC
AC Q99IB6; PRT; 3033 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptidein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Wakita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient.";
RL J. Med. Virol. 64:334-339(2001).
DR EMBL; AB047641; BAB32874.1; -; Genomic_RNA.
DR HSP; Q8JY81; 1CW.
DR SMR; Q99IB6; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
```

```
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol DS ps.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polypeptidein.
SQ SEQUENCE 3033 AA; 328780 MW; 9AB16077CF46517B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 825
Q99IB7_9HEPC
AC Q99IB7; PRT; 3033 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptidein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Wakita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient.";
RL J. Med. Virol. 64:334-339(2001).
DR EMBL; AB047640; BAB32873.1; -; Genomic_RNA.
DR HSP; Q8JY81; 1CW.
DR SMR; Q99IB7; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
```



```
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329035 MW; 74BD3B786F7D7D8 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1700 PDKEVLY 1706

RESULT 826
Q991B8_9HEPC
ID Q991B8_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q991B8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21316767; PubMed=11424123; DOI=10.1002/jmv.1055;
RA Kato T., Furusaka A., Miyamoto M., Date T., Yasui K., Hiramoto J.,
RA Nagayama K., Tanaka T., Wakita T.;
RT "Sequence analysis of hepatitis C virus isolated from a fulminant
RT hepatitis patient.";
RL J. Med. Virol. 64:334-339(2001).
DR EMBL; AB047639; BAB32872.1; -; Genomic_RNA.
DR HSSP; Q8JY51; 1CWX.
DR SMR; Q991B8; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0008023; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0003723; F:serine-type peptidase activity; IEA.
DR GO; GO:0003968; F:structural molecule activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV env.
DR InterPro; IPR002531; HCV NS1.
```

```
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR007094; RNA pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 328825 MW; 7B5BEA79702AF95B CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1700 PDKEVLY 1706

RESULT 827
Q91ZA2_9HEPC
ID Q91ZA2_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q91ZA2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD2a-7;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238485; AAF59944.1; -; Genomic_RNA.
DR HSSP; Q8JY51; 1CWX.
DR SMR; Q91ZA2; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
```

```
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329532 MW; BE13557BPABAIE87 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 828
Q91Z44_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q91Z44;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD2a-4;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238483; AAF59942.1; -; Genomic_RNA.
DR HSP; Q8YI81; LCWX.
DR SMK; Q91Z44; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; P:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
```

```
DR InterPro; IPR002166; HCV_RDRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329421 MW; 5EAABCD220F4D5CB CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 829
Q91ZAS_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q91ZAS;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MD2a-2;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF238482; AAF59941.1; -; Genomic_RNA.
DR HSP; Q8YI81; LCWX.
DR SMK; Q91ZAS; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; P:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RDRP.
```

DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329099 MW; 74E61003AEC7816A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
|||||
Db 1700 PDKEVLY 1706

RESULT 830
Q91ZAG_9HEPC
ID Q91ZAG_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q91ZAG;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MDa-1;
RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
RA Sato C.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF238481; AAF59940.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q91ZAG; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RDRP.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.

DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329445 MW; 9D0F544E06CA2B28 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
|||||
Db 1700 PDKEVLY 1706

RESULT 831
Q909A7_9HEPC
ID Q909A7_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q909A7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21361470; PubMed=11469731; DOI=10.1002/jmv.1073;
RA Kurihara C., Ishiyama N., Nishiyama Y., Fukushi S., Kageyama T.,
RA Katayama K., Miura S.;
RT "Molecular characterization of hepatitis C virus genotype 2a from the
RT entire sequences of four isolates";
RL J. Med. Virol. 64:466-475(2001).
DR EMBL; AF169005; AAF25613.1; -; Genomic_RNA.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q909A7; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006350; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RDRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.

DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00998; RDRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 KW POLYPROTEIN.
 SQ SEQUENCE 3033 AA; 329227 MW; 21492388CA0D5D8C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
 |||||
 DB 1700 PDKEVLY 1706

RESULT 832
 Q9Q9A8_9HEPC
 ID Q9Q9A8_9HEPC PRELIMINARY; PRT; 3033 AA.
 AC Q9Q9A8;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21361470; PubMed=11468731; DOI=10.1002/jmv.1073;
 RA Kurihara C., Ishiyama N., Nishiyama Y., Fukushi S., Kageyama T.,
 RA Katayama K., Miura S.;
 RT "Molecular characterization of hepatitis C virus genotype 2a from the
 RL J. Med. Virol. 64:466-475(2001).
 DR EMBL; AF169004; AAP25612.1; -; Genomic_RNA.
 DR HSP; Q8YIS1; ICWX.
 DR SMR; Q9Q9A8; 1033-1661, 2443-3003.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0016831; F:carboxy-lyase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006520; P:amino acid metabolism; IEA.
 DR GO; GO:0006508; P:protein catabolism; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR00745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.
 DR InterPro; IPR004109; Peptidase_S29.

DR InterPro; IPR002518; Pept_U39 HCV NS2.
 DR InterPro; IPR002129; Pyridoxal deC.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR Pfam; PF01543; HCV_capsid; 1.
 DR Pfam; PF01542; HCV_core; 1.
 DR Pfam; PF01539; HCV_env; 1.
 DR Pfam; PF01560; HCV_NS1; 1.
 DR Pfam; PF01538; HCV_NS2; 1.
 DR Pfam; PF02907; HCV_NS3; 1.
 DR Pfam; PF01006; HCV_NS4a; 1.
 DR Pfam; PF01001; HCV_NS4b; 1.
 DR Pfam; PF01506; HCV_NS5a; 1.
 DR Pfam; PF00998; RDRP_3; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
 DR POLYPROTEIN.
 SQ SEQUENCE 3033 AA; 329175 MW; 74932A91979E78A5 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
 |||||
 DB 1700 PDKEVLY 1706

RESULT 833
 Q9Q9B0_9HEPC
 ID Q9Q9B0_9HEPC PRELIMINARY; PRT; 3033 AA.
 AC Q9Q9B0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21361470; PubMed=11468731; DOI=10.1002/jmv.1073;
 RA Kurihara C., Ishiyama N., Nishiyama Y., Fukushi S., Kageyama T.,
 RA Katayama K., Miura S.;
 RT "Molecular characterization of hepatitis C virus genotype 2a from the
 RL J. Med. Virol. 64:466-475(2001).
 DR EMBL; AF169002; AAP25610.1; -; Genomic_RNA.
 DR HSP; Q8YIS1; ICWX.
 DR SMR; Q9Q9B0; 1033-1661, 2443-3008.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; P:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR011545; DEAD/DEAH N.
 DR InterPro; IPR002522; HCV capsid.
 DR InterPro; IPR002519; HCV env.
 DR InterPro; IPR002531; HCV NS1.
 DR InterPro; IPR00745; HCV NS4a.
 DR InterPro; IPR001490; HCV NS4b.
 DR InterPro; IPR002868; HCV NS5a.
 DR InterPro; IPR002166; HCV RdRP.

```

DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329180 MW; 6E3F51PD3A8A4C24 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred.No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 834
Q9QF35_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q9QF35;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HC-J6CH;
RX MEDLINE=99420396; PubMed=10489358; DOI=10.1006/viro.1999.9889;
RA Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
RT "Hepatitis C virus: an infectious molecular clone of a second major
RT genotype (2a) and lack of viability of intertypic 1a and 2a
RT chimeras.";
RL Virology 262:250-263 (1999).
DR EMBL; AF177036; AAF01178.1; -; Genomic_RNA.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q9QF35; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RGRP.

DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329264 MW; 6D994082FF885C7 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred.No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 835
Q9QAX1_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q9QAX1;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OC NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Genotype 2; TISSUE=Serum;
RX MEDLINE=20329282; PubMed=10872881; DOI=10.1023/A:1008182901274;
RA Sanokhvalov E.I., Hijiata M., Gylka R.I., Lvov D.K., Mishiro S.;
RT "Full-genome nucleotide sequence of a hepatitis C virus variant
RT (isolate name VAR96) representing a new subtype within the genotype 2
RT (arbitrarily 2k).";
RL Virus Genes 20:183-187 (2000).
DR EMBL; AB031663; BAA88057.1; -; Genomic_RNA.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q9QAX1; 1033-1661, 2443-3008.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RGRP.

```

```
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
DR CHAIN 1 191 core protein.
DR CHAIN 192 383 E1 protein.
DR CHAIN 384 733 E2/NS1 protein.
DR CHAIN 734 1010 NS2 protein.
DR CHAIN 1011 1661 NS3 protein.
DR CHAIN 1662 1715 NS4A protein.
DR CHAIN 1716 2017 NS4B protein.
DR CHAIN 2018 2442 NS5A protein.
DR CHAIN 2443 3033 NS5B protein.
DR CHAIN 3033 329860 MW; 5245F9E0A46A7E50 CRC64;
SQ SEQUENCE 3033 AA; 329860 MW; 5245F9E0A46A7E50 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 836
Q7T7H9_9HEPC
ID Q7T7H9_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q7T7H9;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
RA Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232743; AAP55698.1; -; mRNA.
DR HSP; Q8JY81; 1CW.
DR SMR; Q7T7H9; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
```

```
DR InterPro; IPR001410; DEAD
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
DR Polyprotein.
DR CHAIN 1 191 core protein.
DR CHAIN 192 383 E1 protein.
DR CHAIN 384 733 E2/NS1 protein.
DR CHAIN 734 1010 NS2 protein.
DR CHAIN 1011 1661 NS3 protein.
DR CHAIN 1662 1715 NS4A protein.
DR CHAIN 1716 2017 NS4B protein.
DR CHAIN 2018 2442 NS5A protein.
DR CHAIN 2443 3033 NS5B protein.
DR CHAIN 3033 329860 MW; 20CC77C4F89C2AE3 CRC64;
SQ SEQUENCE 3033 AA; 330475 MW; 20CC77C4F89C2AE3 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 837
Q7T7I0_9HEPC
ID Q7T7I0_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q7T7I0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
RA Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232742; AAP55697.1; -; mRNA.
DR HSP; Q8JY81; 1CW.
DR SMR; Q7T7I0; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
```

```
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_P5.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 330368 MW; 46C9F5030F039B12 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1700 PDKEVLY 1706

RESULT 838
Q77T15_9HEPC
ID Q77T15_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q77T15.
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
RA Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232737; AAP55692.1; -; mRNA.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q77T15; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
```

```
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_P5.
DR InterPro; IPR007094; RNA_pol_P5vir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP 3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 330201 MW; 79379F5E329FB997 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1700 PDKEVLY 1706

RESULT 839
Q77T16_9HEPC
ID Q77T16_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q77T16.
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_taxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
RA Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232736; AAP55691.1; -; mRNA.
DR HSP; Q8JYS1; 1CW.
DR SMR; Q77T16; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV capsid.
```

DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01538; HCV NS1; 1.
DR Pfam; PF02907; HCV NS2; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329956 MW; D58C8C087216598C CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 PDKEVLY 56
| | | | |
DB 1700 PDKEVLY 1706

RESULT 840
Q7T719_9HEPC
ID Q7T719_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q7T719_2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232733; AAP55668.1; -; mRNA.
DR HSP; Q8JY81; ICWX.
DR SMR; Q7T719; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; P:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV env.

DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept U39 HCV NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01538; HCV NS1; 1.
DR Pfam; PF02907; HCV NS2; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RdRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 330167 MW; 919ACA2A01686176 CRC64;
Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 60 DEMEECS 66
| | | | |
DB 1710 DEMEECS 1716
RESULT 841
Q7T7J0_9HEPC
ID Q7T7J0_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q7T7J0_2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Nagayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232732; AAP55687.1; -; mRNA.
DR HSP; Q8JY81; ICWX.
DR SMR; Q7T7J0; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0008026; P:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; P:RNA binding; IEA.
DR GO; GO:0003968; P:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; P:serine-type peptidase activity; IEA.
DR GO; GO:0005198; P:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.


```
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 329951 MW; DCDAB491E084582A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred.No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 DMEBES 66
Db 1710 DMEBES 1716

RESULT 842
Q77J1_9HEPC
ID Q77J1_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q77J1_9HEPC PRELIMINARY; PRT; 3033 AA.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed-15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Negayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
RA Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232731; AAP55686.1; -, mRNA.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q77J1; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008024; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR002531; HCV NS1.
```

```
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV NS5a.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00998; RDRP_3; 1.
DR SMART; SM00487; DEXDC; 1.
KW Polyprotein.
SQ SEQUENCE 3033 AA; 330861 MW; C6673E5964AEC019 CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3033;
Best Local Similarity 100.0%; Pred.No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 50 PDKEVLY 56
Db 1700 PDKEVLY 1706

RESULT 843
Q77J2_9HEPC
ID Q77J2_9HEPC PRELIMINARY; PRT; 3033 AA.
AC Q77J2_9HEPC PRELIMINARY; PRT; 3033 AA.
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed-15850465; DOI=10.1111/j.1365-2893.2005.00511.x;
RA Tanabe Y., Negayama K., Enomoto N., Izumi N., Tazawa J., Kurosaki M.,
RA Sakamoto N., Sato C., Watanabe M.;
RT "Characteristic sequence changes of hepatitis C virus genotype 2b
RT associated with sustained biochemical response to IFN therapy.";
RL J. Viral Hepat. 12:251-261(2005).
DR EMBL; AY232731; AAP55686.1; -, mRNA.
DR HSP; Q8JYS1; ICWX.
DR SMR; Q77J2; 1033-1661, 2443-2972.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008024; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH N.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
```

DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RDRP.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV_capsid; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; RDRP_3; 1.
 DR SMART: SM00487; DEXDc; 1.
 KW Polyprotein.
 SQ SEQUENCE 3033 AA; 330641 MW; 593B6BD2358AF44E CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
 |||||
 Db 1700 PDKEVLY 1706

RESULT 844
 Q91ZA3_9HEPC
 ID Q91ZA3_9HEPC PRELIMINARY; PRT; 3033 AA.
 AC Q91ZA3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RC STRAIN-MB2a-5;
 RA Itakura J., Nagayama K., Enomoto N., Kurosaki M., Watanabe H.,
 Sato C.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=93197128; PubMed=8383835;
 RA Sarashina T., Sakurai T., Watanabe Y., Kashima K., Suzuki T.,
 Chiba J., Kita Y., Horiuchi T., Saito I., Miyamura T.;
 RT "Nucleotide sequence of the hepatitis C virus genome from a patient
 negative for anti-HCV by the first generation antibody assay."
 RL Nucleic Acids Res. 21:1037-1037(1993).
 DR EMBL: AF238484; AAF59943.1; -; Genomic_RNA.
 DR FNR; S35631; S35631.
 DR HSP; Q8JY81; 1CW.
 DR SMR; Q91ZA3; 1033-1661, 2443-2972.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; F:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.

DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.
 DR InterPro: IPR000745; HCV_NS4a.
 DR InterPro: IPR001490; HCV_NS4b.
 DR InterPro: IPR002868; HCV_NS5a.
 DR InterPro: IPR002166; HCV_RDRP.
 DR InterPro: IPR002518; Pept U39 HCV NS2.
 DR InterPro: IPR004109; Peptidase S29.
 DR InterPro: IPR007095; RNA_pol_DS_PS.
 DR InterPro: IPR007094; RNA_pol_PSVir.
 DR Pfam: PF01543; HCV_capsid; 1.
 DR Pfam: PF01542; HCV_core; 1.
 DR Pfam: PF01539; HCV_env; 1.
 DR Pfam: PF01560; HCV_NS1; 1.
 DR Pfam: PF01538; HCV_NS2; 1.
 DR Pfam: PF02907; HCV_NS3; 1.
 DR Pfam: PF01006; HCV_NS4a; 1.
 DR Pfam: PF01001; HCV_NS4b; 1.
 DR Pfam: PF01506; HCV_NS5a; 1.
 DR Pfam: PF00998; RDRP_3; 1.
 DR SMART: SM00487; DEXDc; 1.
 KW Polyprotein.
 SQ SEQUENCE 3033 AA; 329519 MW; E016D0E175644593 CRC64;
 Query Match 5.9%; Score 7; DB 2; Length 3033;
 Best Local Similarity 100.0%; Pred. No. 2.5e+03;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
 |||||
 Db 1700 PDKEVLY 1706

RESULT 845
 Q5XNK5_9HEPC
 ID Q5XNK5_9HEPC PRELIMINARY; PRT; 3034 AA.
 AC Q5XNK5;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Polyprotein.
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Amako Y., Kohara K., Kohara M.;
 RT "HCR24 is a hepatitis C virus genotype 2a.";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY746460; AAU89634.1; -; Genomic_RNA.
 DR SMR; Q5XNK5; 1034-1662, 2444-2973.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0019031; C:viral envelope; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
 DR GO; GO:0003723; F:RNA binding; IEA.
 DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
 DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
 DR GO; GO:0006350; F:transcription; IEA.
 DR GO; GO:0019079; P:viral genome replication; IEA.
 DR GO; GO:0019087; P:viral transformation; IEA.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR011545; DEAD/DEAH_N.
 DR InterPro: IPR002522; HCV_capsid.
 DR InterPro: IPR002521; HCV_core.
 DR InterPro: IPR002519; HCV_env.
 DR InterPro: IPR002531; HCV_NS1.

```
DR InterPro; IPR000745; HCV_NS4a.
DR InterPro; IPR001490; HCV_NS4b.
DR InterPro; IPR002868; HCV_NS5a.
DR InterPro; IPR002166; HCV_RGRP.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Pfam; PF01539; HCV_env; 1.
DR Pfam; PF01560; HCV_NS1; 1.
DR Pfam; PF01538; HCV_NS2; 1.
DR Pfam; PF02907; HCV_NS3; 1.
DR Pfam; PF01006; HCV_NS4a; 1.
DR Pfam; PF01001; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS4b; 1.
DR Pfam; PF01506; HCV_NS5a; 1.
DR Pfam; PF00998; RGRP_3; 1.
DR SMART; SM00487; DEXDc; 1.
KW Polyprotein.
SQ SEQUENCE 3034 AA; 329052 MW; 00F69A05457AA94A CRC64;

Query Match 5.9%; Score 7; DB 2; Length 3034;
Best Local Similarity 100.0%; Pred. No. 2.5e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 PDKEVLY 56
DB 1701 PDKEVLY 1707

RESULT 846
Q68749 SHEPC
ID Q68749_9HEPC PRELIMINARY; PRT; 3037 AA.
AC Q68749.
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; asRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=96215652; PubMed=8645105;
RA Nakao H., Okamoto H., Tokita H., Inoue T., Iizuka H., Pozzato G.,
RA Mihiro S.;
RT "Full-length genomic sequence of a hepatitis C virus genotype 2c
RT isolate (BEBE1) and the 2c-specific PCR primers.";
RL Arch. Virol. 141:701-704(1996).
DR EMBL; D50409; BAA08911.1; -; Genomic_RNA.
DR HSSP; Q8JYS1; 1CWK.
DR SMR; Q68749; 1033-1661, 2447-3012.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003969; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR011545; DEAD/DEAH_N.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR InterPro; IPR002519; HCV_env.
DR InterPro; IPR002531; HCV_NS1.
DR InterPro; IPR000745; HCV_NS4a.

Query Match 5.9%; Score 7; DB 2; Length 5183;
Best Local Similarity 100.0%; Pred. No. 3.8e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 3332 VLAALAA 3338

RESULT 848
Q57457 HUMAN
ID Q57457_HUMAN PRELIMINARY; PRT; 5183 AA.
AC Q57457.
DT 01-FEB-2005 (TRENBLrel. 29, Created)
```

DT	01-FEB-2005 (TrEMBLrel. 29, Last sequence update)	
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)	
DE	Retinoblastoma-associated factor 600 (RBAF600).	
GN	Name=RP5-1126H10.1, ORFNames=RP5-1126H10.1-001;	
OS	Homo sapiens (human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Suarchontoglires; Primates; Catarrhini; Hominidae;	
OC	Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Brown A.;	
RL	Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.	
RN	[2]	
RA	Sahra H.;	
RL	Submitted (MAY-2005) to the EMBL/GenBank/DDBJ databases.	
DR	EMBL; AL357564; CAI20974.1; -; Genomic DNA.	
DR	EMBL; AL317127; CAI19272.1; -; Genomic DNA.	
DR	EMBL; AL317127; CAI20974.1; JOINED; Genomic DNA.	
DR	EMBL; AL357564; CAI19272.1; JOINED; Genomic DNA.	
DR	GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.	
DR	GO; GO:0006512; P:ubiquitin cycle; IEA.	
DR	InterPro; IPR001920; Asp/Glu_rac.	
DR	InterPro; IPR007087; Znf C2H2.	
DR	InterPro; IPR003126; Znf Nrecognin.	
DR	Pfam; PF02207; zf-UBR1; 1.	
DR	SMART; SM00396; Znf UBR1; 1.	
DR	PROSITE; PS00923; ASP_GLU_RACEMASE_1; UNKNOWN_1.	
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.	
SQ	SEQUENCE 5183 AA; 573835 MW; 5F6DD7B565827609 CRC64;	
Query Match	5.9%; Score 7; DB 2; Length 5183;	
Best Local Similarity	100.0%; Pred. No. 3.8e+03;	
Matches	7; Conservative 0; Mismatches 0; Indels 0; Gaps	
Qy	20 VLAALAA 26	
Db	3332 VLAALAA 3338	
RESULT 849		
ID	Q93NX8 9ACTO PRELIMINARY; PRT; 5644 AA.	
AC	Q93NX8;	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)	
DE	AmphJ.	
GN	Name=amphJ;	
OS	Streptomyces nodosus.	
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC	Streptomycineae; Streptomycetaceae; Streptomyces.	
OX	NCBI_TaxID=40318;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RX	MEDLINE=21344785; PubMed=11451671; DOI=10.1016/S1074-5521(01)00046-1;	
RA	Caffrey P., Lynch S., Flood E., Finnan S., Olinyk M.;	
RT	"Amphoteric biosynthesis in Streptomyces nodosus: deductions from	
RL	analysis of polyketide synthase and late genes.";	
RL	Chem. Biol. 8:713-723(2001).	
DR	EMBL; AF357202; AAK73502.1; -; Genomic DNA.	
DR	HSP; PF6203; 1PQW	
DR	GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.	
DR	GO; GO:0048037; F:cofactor binding; IEA.	
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.	
DR	GO; GO:0016740; F:transferase activity; IEA.	
DR	GO; GO:0008270; F:zinc ion binding; IEA.	
DR	GO; GO:0006633; P:fatty acid biosynthesis; IEA.	
DR	GO; GO:0008152; P:metabolism; IEA.	
DR	InterPro; IPR009081; ACP_like.	
DR	InterPro; IPR001227; Ac transferase.	
DR	InterPro; IPR002198; ADH short.	

```
DR InterPro; IPR003089; AB_hydrolase.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR001597; Beta_elim_lyase.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR006163; Pheppanteth_bind.
DR InterPro; IPR006162; Ppanteth_S.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00561; Abhydrolase.1; 1.
DR Pfam; PF00106; adh_short; 3.
DR Pfam; PF02122; Beta_elim_lyase; 1.
DR Pfam; PF00109; ketoacyl-synt; 4.
DR Pfam; PF02801; ketoacyl-synt_C; 4.
DR Pfam; PF00550; PP-binding; 8.
DR PRINTS; PR00111; ABHYDROLASE.
DR PRODOM; PD005927; Beta_elim_lyase; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 8.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN 3.
SQ SEQUENCE 7349 AA; 775626 MW; 07FBB7863D66811B CRC64;

Query Match 5.1%; Score 7; DB 2; Length 7349;
Best Local Similarity 100.0%; Pred. No. 5.1e+03;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 VLAALAA 26
DB 3463 VLAALAA 3469

RESULT 851
Q7M2Q1_MACFA
ID Q7M2Q1_MACFA PRELIMINARY; PRT; 18 AA.
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DE Carboxylesterase (EC 3.1.1.1) MK1, microsomal (Fragment).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopitheciidae; Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=90179180; PubMed=2310190;
RA Hosokawa M., Maki T., Satoh T.;
RT "Characterization of molecular species of liver microsomal
RL Arch. Biochem. Biophys. 277:219-227(1990).
DR PIR; S09026; S09026.
DR GO; GO:0004091; F:carboxylesterase activity; IEA.
FT NON TER 1 1
FT NON TER 18 18
SQ SEQUENCE 18 AA; 1779 MW; 14B0BDAE086D25BB CRC64;

Query Match 5.1%; Score 6; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGKVLG 88
DB 12 KGKVLG 17

RESULT 852
Q84274_HPV25
ID Q84274_HPV25 PRELIMINARY; PRT; 19 AA.
AC Q84274;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

DR InterPro; IPR003089; AB_hydrolase.
DR InterPro; IPR009081; ACP_like.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR001597; Beta_elim_lyase.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR001601; Methyltransf.
DR InterPro; IPR006163; Pheppanteth_bind.
DR InterPro; IPR006162; Ppanteth_S.
DR InterPro; IPR000051; SAM_bind.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00561; Abhydrolase.1; 1.
DR Pfam; PF00106; adh_short; 3.
DR Pfam; PF02122; Beta_elim_lyase; 1.
DR Pfam; PF00109; ketoacyl-synt; 4.
DR Pfam; PF02801; ketoacyl-synt_C; 4.
DR Pfam; PF00550; PP-binding; 8.
DR PRINTS; PR00111; ABHYDROLASE.
DR PRODOM; PD005927; Beta_elim_lyase; 1.
DR PROSITE; PS00075; ACP_DOMAIN; 8.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 4.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN 3.
SQ SEQUENCE 7349 AA; 775626 MW; 07FBB7863D66811B CRC64;

Query Match 5.1%; Score 6; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 EQAQVI 78
DB 14 EQAQVI 19

RESULT 853
Q9QUX7_9MURI
ID Q9QUX7_9MURI PRELIMINARY; PRT; 20 AA.
AC Q9QUX7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Carboxylesterase isozyme (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
RP PROTEIN SEQUENCE.
RX MEDLINE=96170127; PubMed=8597091; DOI=10.1016/0378-4274(95)03493-5;
RA Satoh T., Hosokawa M.;
RT "Molecular aspects of carboxylesterase isoforms in comparison with
RL Toxicol. Lett. 82:439-445(1995).
SQ SEQUENCE 20 AA; 2133 MW; 435160FFA80E086D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 KGKVLG 88
DB 12 KGKVLG 17

RESULT 854
Q9QUX6_9MURI
ID Q9QUX6_9MURI PRELIMINARY; PRT; 22 AA.
AC Q9QUX6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
DE Carboxylesterase isozyme (Fragment).
OS Rattus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10118;
RN [1]
```

```
RP PROTEIN SEQUENCE.
RX MEDLINE=96170127; PubMed=8597091; DOI=10.1016/0378-4274(95)03493-5;
RA Satoh T., Hosokawa M.;
RT "Molecular aspects of carboxylesterase isoforms in comparison with
  Other esterases.";
RL Toxicol. Lett. 82:439-445(1995).
SQ SEQUENCE 22 AA; 2286 MW; 6AE5635160E2D03E CRC64;

Query Match          5.1%; Score 6; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 KGKVLG 88
Db 12 KGKVLG 17

RESULT 855
Q4RAB0_TETNG
ID Q4RAB0_TETNG PRELIMINARY; PRT; 29 AA.
AC Q4RAB0_TETNG
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE Chromosome undetermined SCAF24109, whole genome shotgun sequence.
DE (Fragment).
ORFNames=GSTENG00036872001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OC NCBI_TaxID=99883;
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]
[908]
[909]
[910]
[911]
[912]
[913]
[914]
[915]
[916]
[917]
[918]
[919]
[920]
[921]
[922]
[923]
[924]
[925]
[926]
[927]
[928]
[929]
[930]
[931]
[932]
[933]
[934]
[935]
[936]
[937]
[938]
[939]
[940]
[941]
[942]
[943]
[944]
[945]
[946]
[947]
[948]
[949]
[950]
[951]
[952]
[953]
[954]
[955]
[956]
[957]
[958]
[959]
[960]
[961]
[962]
[963]
[964]
[965]
[966]
[967]
[968]
[969]
[970]
[971]
[972]
[973]
[974]
[975]
[976]
[977]
[978]
[979]
[980]
[981]
[982]
[983]
[984]
[985]
[986]
[987]
[988]
[989]
[990]
[991]
[992]
[993]
[994]
[995]
[996]
[997]
[998]
[999]
[1000]
```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
RN NCBI_TaxID=314565;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=8004;
RA Qian W., Jia Y.-T., Ren S.-X., He Y.-Q., Feng J.-X., Lu L.-F.,
RA Sun Q.-H., Yang G., Tang D.-J., Wu W., Wang L.-F., Jiang B.-L.,
RA Zeng S.-Y., Gu W.-Y., Lu G., Rong L., Tian Y.-C., Yao Z.-J., Fu G.,
RA Chen B.-S., Fang R.-X., Qiang B.-Q., Chen Z., Zhao G.-P., Tang J.-L.,
RA He C.-Z.;
RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; CP000050; AAY48946.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 33 AA; 3707 MW; C92343B90171FD67 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 12 LAALAA 17

RESULT 859
Q8P8K6 XANCP PRELIMINARY; PRT; 33 AA.
AC Q8P8K6
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein XCC2234.
GN OrderedLocusNames=XCC2234;
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 33913 / NCPPB 528;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AB012331; AAM41514.1; -; Genomic_DNA.
DR InterPro; IPR011724; Cvd oper_ybgt.
DR TIGRFAMs; TIGR02106; cvd oper_ybgt; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 3767 MW; C92343B90165AC22 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 12 LAALAA 17

```

```

RESULT 860
Q7RSC3 PLAYO PRELIMINARY; PRT; 36 AA.
AC Q7RSC3;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=PY00437;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=17XNL;
RX MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feidblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Praser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC Preliminary data.
DR EMBL; AB01000120; EAA15834.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 36 AA; 3523 MW; 3C88807DAA09526D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 6.9e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
Db 18 GKVLGL 23

RESULT 861
Q8PK38 XANAC PRELIMINARY; PRT; 36 AA.
AC Q8PK38
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein XAC2338.
GN OrderedLocusNames=XAC2338;
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;
[1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=306 / ATCC 13902 / XV 101;
RX MEDLINE=2202145; PubMed=12024217; DOI=10.1038/417459a;
RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A.,
RA Almeida N.F. Jr., Alves L.M.C., do Amaral A.M., Bertolini M.C.,
RA Camargo L.E.A., Camarotte G., Cannavan F., Cardozo J., Chambergo F.,
RA Ciapina L.P., Ciccarelli R.M.B., Coutinho L.L., Cursino-Santos J.R.,
RA El-Dorry H., Faria J.B., Ferreira A.J.S., Ferreira R.C.C.,
RA Ferro M.I.T., Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Teai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463(2002).
DR EMBL; AB012331; AAM41514.1; -; Genomic_DNA.
DR InterPro; IPR011724; Cvd oper_ybgt.
DR TIGRFAMs; TIGR02106; cvd oper_ybgt; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 33 AA; 3767 MW; C92343B90165AC22 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 6.4e+02; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 12 LAALAA 17

```

```
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RT "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities.";
RL Nature 417:459-463(2002).
DR EMBL; AE011871; AM37191.1; -; Genomic_DNA.
DR InterPro; IPR011724; Cvd_oper_ybgt.
DR TIGRFAMs; TIGR02106; cvd_oper_ybgt; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 36 AA; 4070 MW; DE50D2A3CC56468B0 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 12 LAALAA 17
|||||

RESULT 862
Q81247_9HEPC
ID Q81247_9HEPC PRELIMINARY; PRT; 41 AA.
AC Q81247;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NS4 (Fragment).
GN Name=HCV-C10;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92068204; PubMed=1720309;
RA Kato N., Ootsuyama Y., Ohkoshi S., Nakazawa T., Mori S., Hijikata M.,
RA Shimotohno K.;
RT "Distribution of plural HCV types in Japan.";
RL Biochem. Biophys. Res. Commun. 181:279-285(1991).
DR EMBL; D10565; BAA01421.1; -; Genomic_RNA.
DR PIR; PQ0562; PQ0562.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR NON_TER 1
FT NON_TER 41
SQ SEQUENCE 41 AA; 4131 MW; F135A8C5D2D8CEC1 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9
|||||

RESULT 863
Q81248_9HEPC
ID Q81248_9HEPC PRELIMINARY; PRT; 41 AA.
AC Q81248;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NS4 (Fragment).
GN Name=HCV-C11;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
```

```
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92068204; PubMed=1720309;
RA Kato N., Ootsuyama Y., Ohkoshi S., Nakazawa T., Mori S., Hijikata M.,
RA Shimotohno K.;
RT "Distribution of plural HCV types in Japan.";
RL Biochem. Biophys. Res. Commun. 181:279-285(1991).
DR EMBL; D10566; BAA01422.1; -; Genomic_RNA.
DR PIR; PQ0563; PQ0563.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR NON_TER 1
FT NON_TER 41
SQ SEQUENCE 41 AA; 4120 MW; 1C55A8D496D8C1F5 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9
|||||

RESULT 864
Q81249_9HEPC
ID Q81249_9HEPC PRELIMINARY; PRT; 41 AA.
AC Q81249;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NS4 (Fragment).
GN Name=HCV-C13;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92068204; PubMed=1720309;
RA Kato N., Ootsuyama Y., Ohkoshi S., Nakazawa T., Mori S., Hijikata M.,
RA Shimotohno K.;
RT "Distribution of plural HCV types in Japan.";
RL Biochem. Biophys. Res. Commun. 181:279-285(1991).
DR EMBL; D10568; BAA01424.1; -; Genomic_RNA.
DR PIR; PQ0565; PQ0565.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
DR NON_TER 1
FT NON_TER 41
SQ SEQUENCE 41 AA; 4148 MW; 1C55B01AD3D25F25 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9
|||||

RESULT 865
Q81251_9HEPC
ID Q81251_9HEPC PRELIMINARY; PRT; 41 AA.
AC Q81251;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE NS4 (Fragment).
GN Name=HCV-C4;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
```



```

OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92068204; PubMed=1720309;
RA Kato N., Ootsuyama Y., Ohkoshi S., Nakazawa T., Mori S., Hijikata M.,
RA Shimotohno K.;
RT "Distribution of plural HCV types in Japan.";
RL Biochem. Biophys. Res. Commun. 181:279-285(1991).
DR EMBL; D10567; BAA01423.1; -; Genomic_RNA.
DR PIR; PQ0564; PQ0564.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 1
FT NON_TER 41
SQ SEQUENCE 41 AA; 4152 MW; 6AF4401AD3C42PF81 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9

RESULT 866
Q81252_9HEPC PRELIMINARY; PRT; 41 AA.
AC Q81252_
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE NS4 (Fragment).
GN Name=HCV-C8;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92068204; PubMed=1720309;
RA Kato N., Ootsuyama Y., Ohkoshi S., Nakazawa T., Mori S., Hijikata M.,
RA Shimotohno K.;
RT "Distribution of plural HCV types in Japan.";
RL Biochem. Biophys. Res. Commun. 181:279-285(1991).
DR EMBL; D10563; BAA01419.1; -; Genomic_RNA.
DR PIR; PQ0560; PQ0560.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 1
FT NON_TER 41
SQ SEQUENCE 41 AA; 4133 MW; 1C55A8C5D2C062F5 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9

RESULT 867
Q81253_9HEPC PRELIMINARY; PRT; 41 AA.
AC Q81253_
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE NS4 (Fragment).
GN Name=HCV-C9;
OS Hepatitis C virus.

```

```

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=92068204; PubMed=1720309;
RA Kato N., Ootsuyama Y., Ohkoshi S., Nakazawa T., Mori S., Hijikata M.,
RA Shimotohno K.;
RT "Distribution of plural HCV types in Japan.";
RL Biochem. Biophys. Res. Commun. 181:279-285(1991).
DR EMBL; D10564; BAA01420.1; -; Genomic_RNA.
DR PIR; PQ0561; PQ0561.
DR InterPro; IPR000745; HCV NS4a.
DR Pfam; PF01006; HCV_NS4a; 1.
FT NON_TER 1
FT NON_TER 41
SQ SEQUENCE 41 AA; 4170 MW; 773804D5C3C162E4 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 GGVLA 23
Db 4 GGVLA 9

RESULT 868
Q83Q33_BURPS PRELIMINARY; PRT; 42 AA.
ID Q83Q33_BURPS
AC Q83Q33_
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Putative lipoprotein.
GN OrderedlocusNames=BPSL3060;
OS Burkholderia pseudomallei (pseudomonas pseudomallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia; pseudomallei group.
OX NCBI_TaxID=28450;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K96243;
RX PubMed=15377794; DOI=10.1073/pnas.0403302101;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdeno-Tarraga A.-M.,
RA Atkins T.C., Crossman L.C., Pitt T., Churcher C., Mungall K.L.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosssett B., Davis P., DeShazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitsch E., Rutherford K., Sanders M., Simmonds M.,
RA Songsvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RL Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH37071.1; -; Genomic_DNA.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 42 AA; 4663 MW; 6F65BF9B3FE59A62 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 6 VLAALA 11

RESULT 869
Q62DD7_BURMA PRELIMINARY; PRT; 42 AA.
ID Q62DD7_BURMA

```

```

AC Q62DD7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocusNames=BMAA0523;
OS Burkholderia mallei (Pseudomonas mallei).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=13373;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ATCC 23344;
RX PubMed=15377793; DOI=10.1073/pnas.0403306101;
RA Nierman W.C., Deshazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T.V., Ulrich R.D., Ronning C.M., Brinkac L.M.,
RA Daugherty S.C., Davidson T.D., DeBoy R.T., Dimitrov G., Dodson R.J.,
RA Durkin A.S., Gwinn M.L., Haft D.H., Khouri H.M., Kolonay J.F.,
RA Madupu R., Mohammed Y., Nelson W.C., Radune D., Romero C.M.,
RA Sarria S., Selengut J., Shamlin C., Sullivan S.A., White O., Yu Y.,
RA Zafar N., Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14246-14251(2004).
DR EMBL; CP000011; AAU46899.1; -; Genomic_DNA.
DR TIGR; BMAA0523; -;
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 42 AA; 4637 MW; 6F65AA8E3FE59A62 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 7.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 6 VLAALA 11
|||||
|||||

RESULT 870
QSLPR6_SILPO
ID QSLPR6_SILPO PRELIMINARY; PRT; 45 AA.
AC QSLPR6;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SPO2783;
OS Silicibacter pomeroyi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC Rhodobacteraceae; Silicibacter.
OX NCBI_TaxID=89184;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=DSS-3 / ATCC 700808 / DSM 15171;
RX PubMed=15602564; DOI=10.1038/nature031170;
RA Moran M.A., Buchan A., Gonzalez J.M., Heidelberg J.F., Whitman W.B.,
RA Brinkac L.M., Lewis M., Johri S., Weaver B., Pai G., Eisen J.A.,
RA Rahe E., Sheldon W.M., Ye W., Miller T.R., Carlton J., Rasko D.A.,
RA Paulsen I.T., Ren Q., Daugherty S.C., DeBoy R.T., Dodson R.J.,
RA Durkin A.S., Madupu R., Nelson W.C., Sullivan S.A., Rosovitz M.J.,
RA Haft D.H., Selengut J., Ward N.;
RT "Genome sequence of Silicibacter pomeroyi reveals adaptations to the
RT marine environment.";
RL Nature 432:910-913(2004).
DR EMBL; CP000031; AAV96024.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 45 AA; 4555 MW; 4AE183CA99C8FC34 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

```

```

Db 8 LAALAA 13
|||||

RESULT 871
Q8SKD4_ACISU
ID Q8SKD4_ACISU PRELIMINARY; PRT; 45 AA.
AC Q8SKD4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Cytochrome b (Fragment).
OS Acipenser sturio (Atlantic sturgeon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=61674;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC Fopp D., Ciesielski S., Luczynski M.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF468655; AAL76270.1; -; Genomic DNA.
DR EMBL; AY971520; AAX82623.1; -; Genomic DNA.
DR EMBL; AY442324; AAR15152.1; -; Genomic DNA.
DR EMBL; AY971519; AAX82622.1; -; Genomic DNA.
DR SMR; Q8SKD4; 1-45.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_b6_C1.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 45
SQ SEQUENCE 45 AA; 5042 MW; 4E4B6ABF569DA362 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 4 LGGVLA 9
|||||

RESULT 872
Q4RCR3_TETNG
ID Q4RCR3_TETNG PRELIMINARY; PRT; 47 AA.
AC Q4RCR3;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF18241, whole genome shotgun sequence.
DE (Fragment).

```

ORFNames=GSTENG00036908001;
 GN Tetraodon nigroviridis (Green puffer).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jallou O., Aury J.M., Brunet P., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bounoue L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulin J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,
 RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Boeak S.,
 RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weissbach J., Roest Crolius H.,
 RA "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 CC DR EMBL; CA001018241; CAG13820.1; -; Genomic_DNA.
 FT NON_TER 1 1
 FT NON_TER 47 47
 SQ SEQUENCE 47 AA; 5504 MW; 5F754CE86C27B65B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 47;
 Best Local Similarity 100.0%; Pred. No. 8.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 88 GLLQRA 93
 DB 20 GLLQRA 25
 RESULT 873
 Q94VL8 ACIGU PRELIMINARY; PRT; 50 AA.
 ID Q94VL8 ACIGU PRELIMINARY; PRT; 50 AA.
 AC Q94VL8;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser guldenstadti (Caspian sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=7902;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., Doukakis P., DeSalle R.,
 RT "Molecular phylogeny of Acipenseridae: non-monophyly of
 SC Scaphirhynchinae.";
 RL Copeia 2:287-301(2002).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF040809; AAL09666.1; -; Genomic_DNA.
 DR SMR; Q94VL8; 1-50.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom_b6_C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22

DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom_b6_C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22
 RESULT 874
 O63380 ACIME PRELIMINARY; PRT; 50 AA.
 ID O63380 ACIME PRELIMINARY; PRT; 50 AA.
 AC O63380;
 DT 01-AUG-1998 (TREMELrel. 07, Created)
 DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
 DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser medirostris (Green sturgeon).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=7908;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., DeSalle R.,
 RT "Molecular phylogeny of Acipenserinae.";
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF006171; AAC08506.1; -; Genomic_DNA.
 DR SMR; O63380; 1-50.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom_b6_C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22

```
RESULT 875
O63381.ACIBSE PRELIMINARY; PRT; 50 AA.
AC O63381;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Acipenser baerii (Siberian sturgeon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=27689;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
EMBL; AF006172; AAC08507.1; -; Genomic_DNA.
DR SMR; O63381; 1-50.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_b6_C.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 50 50
SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22

RESULT 876
O63392.POLSP PRELIMINARY; PRT; 50 AA.
AC O63392;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Polyodon spathula (North American paddlefish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Polyodontidae;
OC Polyodon.
OX NCBI_TaxID=7913;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22
```

```
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
EMBL; AF006184; AAC08519.1; -; Genomic_DNA.
DR SMR; O63392; 1-47.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_b6_C.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 50 50
SQ SEQUENCE 50 AA; 5708 MW; 73D7DD8E0F23FB2E CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22

RESULT 877
O63376.9ACTI PRELIMINARY; PRT; 50 AA.
AC O63376;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Psephurus gladius (Chinese swordfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Polyodontidae;
OC Psephurus.
OX NCBI_TaxID=61976;
[1]
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
EMBL; AF006167; AAC08502.1; -; Genomic_DNA.
DR SMR; O63376; 1-50.
DR GO; GO:0016021; C:Integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_b6_C.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 50 50
SQ SEQUENCE 50 AA; 5750 MW; 788BFBCB74FFB2B CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
```

Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|
|
|
|
Db 17 LGGVLA 22

RESULT 878
O63385.ACISU
ID O63385 ACISU PRELIMINARY; PRT; 50 AA.
AC O63385;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Name=cytb;
OC Acipenser sturio (Atlantic sturgeon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=61674;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RL "Molecular phylogeny of Acipenserinae";
RM Mol. Phylogenet. Evol. 9:141-155(1998).
CC -I- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
CYTOCHROME c1 and the Rieske protein (By similarity).
DR EMBL; AF006176; AAC08511.1; -; Genomic_DNA.
DR SMR; O63385; 2-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5577 MW; 8D864023BEE3FB36 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|
|
|
|
Db 17 LGGVLA 22

RESULT 879
Q94VM0.ACINA
ID Q94VM0 ACINA PRELIMINARY; PRT; 50 AA.
AC Q94VM0;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Name=cytb;
OC Acipenser naccarii (Adriatic sturgeon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.

KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5733 MW; BF9141FE0F23E599 CRC64;
Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LGGVLA 22
DB 17 LGGVLA 22
RESULT 881
O63377 ACIFR
ID O63377 ACIFR PRELIMINARY; PRT; 50 AA.
AC O63377;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name-cytb;
OS Acipenser ruthenus (Sterlet).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=7906;
RN NUCLEOTIDE SEQUENCE
RP MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF006168; AAC08503.1; -; Genomic_DNA.
DR SMR; O63377; 1-50.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrion; IEA.
DR GO; GO:0046872; P: metal ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5692 MW; 8D815B4E1B66EF76 CRC64;
Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LGGVLA 22
DB 17 LGGVLA 22
RESULT 882
O63387 ACIFR
ID O63387 ACIFR PRELIMINARY; PRT; 50 AA.
AC O63387;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Cytochrome b (Fragment).
GN Name-cytb; fulvescens (lake sturgeon).
OS Acipenser fulvescens (lake sturgeon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=41871;
RN NUCLEOTIDE SEQUENCE
RP MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF006178; AAC08513.1; -; Genomic_DNA.
DR SMR; O63387; 1-50.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrion; IEA.
DR GO; GO:0046872; P: metal ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LGGVLA 22
DB 17 LGGVLA 22
RESULT 883
O63382 ACIFR
ID O63382 ACIFR PRELIMINARY; PRT; 50 AA.
AC O63382;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name-cytb;
OS Acipenser brevirostrum (Shortnose sturgeon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=7907;
RN NUCLEOTIDE SEQUENCE
RP MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF006173; AAC08508.1; -; Genomic_DNA.
DR SMR; O63382; 1-50.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrion; IEA.

DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR005798; Cytochrome b6 C.
 DR Pfam: PF00032; Cytochrome b6 C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 50
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 17 LGGVLA 22

RESULT 884
 Q94VM4 ACIPE
 ID Q94VM4 ACIPE PRELIMINARY; PRT; 50 AA.
 AC Q94VM4;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser persicus (Persian sturgeon).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 CC Cytochrome b (Fragment).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF404802; AAL09659.1; -; Genomic_DNA.
 DR SMR; Q94VM4; 1-50.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR005798; Cytochrome b6 C.
 DR Pfam: PF00032; Cytochrome b6 C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 50
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 17 LGGVLA 22

RESULT 885
 Q94VM2 ACISC
 ID Q94VM2 ACISC PRELIMINARY; PRT; 50 AA.
 AC Q94VM2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser schrenckii (Amur sturgeon).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 CC Cytochrome b (Fragment).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF404805; AAL09662.1; -; Genomic_DNA.
 DR SMR; Q94VM2; 1-50.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR005798; Cytochrome b6 C.
 DR Pfam: PF00032; Cytochrome b6 C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 50
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 17 LGGVLA 22

RESULT 886
 Q94VL9 ACITR
 ID Q94VL9 ACITR PRELIMINARY; PRT; 50 AA.
 AC Q94VL9;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser transmontanus (White sturgeon).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 CC Cytochrome b (Fragment).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF404802; AAL09659.1; -; Genomic_DNA.
 DR SMR; Q94VM4; 1-50.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR005798; Cytochrome b6 C.
 DR Pfam: PF00032; Cytochrome b6 C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 50
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 17 LGGVLA 22

cytochrome c1 and the Rieske protein (By similarity).
 EMBL; AF404808; AAL09665.1; -; Genomic_DNA.
 DR SMR; Q94VL9; 1-50.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C: mitochondrion; IEA.
 DR GO; GO:0046872; P: metal ion binding; IEA.
 DR GO; GO:0016491; F: oxidoreductase activity; IEA.
 DR GO; GO:0006118; P: electron transport; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR Pfam; PF00032; Cytochrom b C; 1.
 DR InterPro; IPR005798; Cytb b6 C.
 DR PROSITE; PS1003; CYTB_CTER; 1.
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KM Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5639 MW; 7C915B47B97CC62B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22
 RESULT 887
 Q94VL6_9ACTI
 ID Q94VL6_9ACTI PRELIMINARY; PRT; 50 AA.
 AC Q94VL6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 GN Scaphirhynchus suttkusi (Alabama sturgeon).
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Scaphirhynchus.
 OX NCBI_TaxID=36179;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., Doukakos P., DeSalle R.;
 RT "Molecular phylogeny of Acipenseridae: non-monophyly of
 Scaphirhynchinae".
 RL Copeia 2:287-301(2002).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 EMBL; AF404813; AAL09670.1; -; Genomic_DNA.
 DR SMR; Q94VL6; 1-50.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C: mitochondrion; IEA.
 DR GO; GO:0046872; P: metal ion binding; IEA.
 DR GO; GO:0016491; F: oxidoreductase activity; IEA.
 DR GO; GO:0006118; P: electron transport; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR005798; Cytb b6 C.
 DR PROSITE; PS1003; CYTB_CTER; 1.
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KM Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5826 MW; E32BF9DC0E23F48E CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22
 RESULT 888
 Q94VM3_ACIBE
 ID Q94VM3_ACIBE PRELIMINARY; PRT; 50 AA.
 AC Q94VM3;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 GN Acipenser baerii baicalensis (Baikal sturgeon).
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=101766;
 RN [1]
 RP NUCLEOTIDE SEQUENCE
 RA Birstein V.J., Doukakos P., DeSalle R.;
 RT "Molecular phylogeny of Acipenseridae: non-monophyly of
 Scaphirhynchinae".
 RL Copeia 2:287-301(2002).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 EMBL; AF404803; AAL09660.1; -; Genomic_DNA.
 DR SMR; Q94VM3; 1-50.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C: mitochondrion; IEA.
 DR GO; GO:0046872; P: metal ion binding; IEA.
 DR GO; GO:0016491; F: oxidoreductase activity; IEA.
 DR GO; GO:0006118; P: electron transport; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR005798; Cytb b6 C.
 DR PROSITE; PS1003; CYTB_CTER; 1.
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KM Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 50 50
 SQ SEQUENCE 50 AA; 5690 MW; 8D815B4E0F3597E6 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 DB 17 LGGVLA 22
 RESULT 889
 O63384_9ACTI
 ID O63384_9ACTI PRELIMINARY; PRT; 50 AA.
 AC O63384;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 GN Huso dauricus.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae; Huso.
 OX NCBI_TaxID=55293;
 RN [1]


```
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF006175; AAC08510.1; -; Genomic_DNA.
DR SMR; O63384; 1-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; P:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5676 MW; 8D815B4E0F23FB36 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db |||||
17 LGGVLA 22

RESULT 890
O63369 ACIMI PRELIMINARY; PRT; 50 AA.
ID O63369
AC O63369
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Acipenser mikadoi (Sakhalin sturgeon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=61966;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF006179; AAC08514.1; -; Genomic_DNA.
DR SMR; O63388; 1-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; P:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5676 MW; 8D815B4E0F23FB36 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db |||||
17 LGGVLA 22

RESULT 891
O63388 ACIGU PRELIMINARY; PRT; 50 AA.
ID O63388
AC O63388
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Acipenser guldenstadti (Caspian sturgeon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=7902;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., Desalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RL Mol. Phylogenet. Evol. 9:141-155(1998).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF006179; AAC08514.1; -; Genomic_DNA.
DR SMR; O63388; 1-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; P:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5676 MW; 8D815B4E0F23FB36 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db |||||
17 LGGVLA 22

RESULT 892
O94VMI ACIMI PRELIMINARY; PRT; 50 AA.
ID O94VMI
AC O94VMI
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
```



```
ID O63383_9ACTI PRELIMINARY; PRT; 50 AA.
AC 063383;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Huso huso (beluga).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae; Huso.
OX NCBI_TaxID=61971;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Birstein V.J., Doukakis P., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Copeia 2:287-301(2002).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
EMBL; AF040804; AAC09661.1; -; Genomic_DNA.
DR EMBL; AF040804; AAC09661.1; -; Genomic_DNA.
DR SMR; 063379; 1-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5676 MW; 8D815B4E0F23FB36 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22

RESULT 896
ID O63379_ACIST PRELIMINARY; PRT; 50 AA.
AC 063379;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Acipenser stellatus (Sevruga).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OX NCBI_TaxID=7903;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
```

```
RA Birstein V.J., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Birstein V.J., Doukakis P., DeSalle R.;
RT "Molecular phylogeny of Acipenseridae: non-monophyly of
Scaphirhynchinae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Copeia 2:287-301(2002).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
EMBL; AF040804; AAC09661.1; -; Genomic_DNA.
DR EMBL; AF040804; AAC09661.1; -; Genomic_DNA.
DR SMR; 063379; 1-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 50
SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22

RESULT 897
ID O63393_9ACTI PRELIMINARY; PRT; 50 AA.
AC 063393;
DT 01-AUG-1998 (TRENBLrel. 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=cytb;
OS Pseudoscaphirhynchus kaufmanni (Amu Darya sturgeon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Pseudoscaphirhynchus.
OX NCBI_TaxID=55290;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
RA Birstein V.J., DeSalle R.;
RT "Molecular phylogeny of Acipenserinae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Birstein V.J., Doukakis P., DeSalle R.;
RT "Molecular phylogeny of Acipenseridae: non-monophyly of
Scaphirhynchinae.";
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Copeia 2:287-301(2002).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
EMBL; AF006185; AAC08520.1; -; Genomic_DNA.
DR EMBL; AF006185; AAC08520.1; -; Genomic_DNA.
DR SMR; 063393; 1-50.
DR GO; GO:0016021; C:integral to membrane; IEA.
```

DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:0046872; F:metal ion binding; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006118; F:electron transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR005798; Cytb_b6_C.
DR Pfam: PF00032; Cytochrom_b6_C1.
DR PROSITE: PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON TER 1 1
FT NON TER 50 50
SQ SEQUENCE 50 AA; 5708 MW; 7C915B4E0F23FB2B CRC64;

Query Match 5.1%; Score 6; DB 2; Length 50;
Best Local Similarity 100.0%; Pred. No. 9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
|
Db 17 LGGVLA 22

RESULT 898

Y14_BPT7
ID Y14_BPT7 STANDARD; PRT; 51 AA.

AC P03791;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Hypothetical gene 1.4 protein.
GN Name=1.4;
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
locations of T7 genetic elements.";
RL J. Mol. Biol. 166:477-535(1983).
RN [2]

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=82078034; PubMed=7310871;
RA Dunn J.J., Studier F.W.;
RT "Nucleotide sequence from the genetic left end of bacteriophage T7 DNA
to the beginning of gene 4.";
RL J. Mol. Biol. 148:303-330(1981).
RN [3]

RP NUCLEOTIDE SEQUENCE OF 1-50.
RX MEDLINE=81053683; PubMed=7001354;
RA Dunn J.J., Studier F.W.;
RT "The transcription termination site at the end of the early region of
bacteriophage T7 DNA.";
RL Nucleic Acids Res. 8:2119-2132(1980).
RN [4]
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----

DR EMBL; V01146; CAA24394.1; -; Genomic DNA.
DR EMBL; V01124; CAA24323.1; -; Genomic DNA.
DR EMBL; V01127; CAA24337.1; -; Genomic DNA.
DR PIR; H43002; Q1BP47.
KW Hypothetical protein.
SQ SEQUENCE 51 AA; 5447 MW; 110DBE006B6F9DB1 CRC64;

Query Match 5.1%; Score 6; DB 1; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|
Db 9 LAALAA 14

RESULT 899

Q38509_BPT7
ID Q38509_BPT7 PRELIMINARY; PRT; 51 AA.

AC Q38509;
DT 01-NOV-1996 (T-EMBLrel. 01, Created)
DT 01-NOV-1996 (T-EMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE ORF 1.4.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=80144900; PubMed=231766;
RA Boothroyd J.C., Hayward R.S.;
RT "New genes and promoters suggested by the DNA sequence near the end of
the coliphage T7 early operon.";
RL Nucleic Acids Res. 7:1931-1943(1979).
DR EMBL; M25443; AAA32565.1; -; Genomic DNA.
SQ SEQUENCE 51 AA; 5479 MW; A60C6D906CE8B868 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|
Db 9 LAALAA 14

RESULT 900

Q6WY53_BPT7
ID Q6WY53_BPT7 PRELIMINARY; PRT; 51 AA.

AC Q6WY53;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Gene 1.4.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22990536; PubMed=14629033; DOI=10.1007/s00239-003-2470-1;
RA Bull J.J., Badgett M.R., Rokyt D., Molineux I.J.;
RT "Experimental evolution yields hundreds of mutations in a functional
viral genome.";
RL J. Mol. Evol. 57:241-248(2003).

DR EMBL; AY264778; AAP34128.1; -; Genomic DNA.
DR EMBL; AY264777; AAP34074.1; -; Genomic DNA.
SQ SEQUENCE 51 AA; 5461 MW; 110DAC906B6F9DB1 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
|
Db 9 LAALAA 14

RESULT 901

OSWYP7 BPT7
ID Q6WYF7_BPT7 PRELIMINARY; PRT; 51 AA.
AC Q6WYF7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Gene 1.4.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22990536; PubMed=14629033; DOI=10.1007/s00239-003-2470-1;
RA Bull J.J., Badgett M.R., Rokya D., Molineux I.J.;
RT "Experimental evolution yields hundreds of mutations in a functional viral genome.";
RL J. Mol. Evol. 57:241-248 (2003).
DR EMBL; AY264776; AAP33973.1; -; Genomic DNA.
DR EMBL; AY264775; AAP33973.1; -; Genomic DNA.
SQ SEQUENCE 51 AA; 5507 MW; 111069A0BA6F9DB1 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 21 LAALAA 26
Db 9 LAALAA 14
|||||

RESULT 902
Q6WYR3_BPT7
ID Q6WYR3_BPT7 PRELIMINARY; PRT; 51 AA.
AC Q6WYR3;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Gene 1.4.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxID=10760;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22990536; PubMed=14629033; DOI=10.1007/s00239-003-2470-1;
RA Bull J.J., Badgett M.R., Rokya D., Molineux I.J.;
RT "Experimental evolution yields hundreds of mutations in a functional viral genome.";
RL J. Mol. Evol. 57:241-248 (2003).
DR EMBL; AY264774; AAP33918.1; -; Genomic DNA.
DR EMBL; AY264774; AAP33918.1; -; Genomic DNA.
SQ SEQUENCE 51 AA; 5447 MW; 110DBE006B6F9DB1 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 21 LAALAA 26
Db 9 LAALAA 14
|||||

RESULT 903
Q4TFP6_TETNG
ID Q4TFP6_TETNG PRELIMINARY; PRT; 52 AA.
AC Q4TFP6;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAF4536, whole genome shotgun sequence. (Fragment).
DE (Fragment).
GN ORFNames=GSTENG00001660001;
OS Tetraodon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Bienoud V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Cruaud C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Parra G., Lardier G., Chappie C., Coutanceau J.P., Gouzy J.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype.";
RL Nature 431:946-957 (2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAS01004536; CAF88376.1; -; Genomic DNA.
FT NON_TER 1 1
FT NON_TER 52 52
SQ SEQUENCE 52 AA; 5990 MW; B10B0EAB92116BC9 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 9.3e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 6; Conservative 0;

Qy 88 GLLQRA 93
Db 25 GLLQRA 30
|||||

RESULT 904
Q4YM45_PLABE
ID Q4YM45_PLABE PRELIMINARY; PRT; 54 AA.
AC Q4YM45;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein (Fragment).
GN ORFNames=PB400501.00.0;
OS Plasmodium berghei.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5821;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA Herriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA Bidwell S.L., Rajandream M.A., Carucci D.J., Yates J.R., Kafatos F.C.,
RA Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT "A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses.";
RL Science 307:82-86 (2005).
CC -!- CAUTION: The sequence shown here is derived from an
CC preliminary data.
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAL01003693; CAI00916.1; -; Genomic DNA.
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 54 54

```
SQ SEQUENCE 54 AA; 6165 MW; 07691C561D0256E3 CRC64;
Query Match 5.1%; Score 6; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 39 HIELOGG 44
Db 35 HIELOGG 40

RESULT 905
Q4NTN0_9DELTA PRELIMINARY; PRT; 55 AA.
AC Q4NTN0_9DELTA PRELIMINARY;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein precursor.
GN ORFNames=AdehdRAFT_2075;
OS Anaeromyxobacter dehalogenans 2CP-C.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacterineae; Myxococcaceae; Anaeromyxobacter.
OX NCBI_TaxID=290397;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=2CP-C;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Anaeromyxobacter
RT dehalogenans 2CP-C.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; AAHD01000017; EAL78960.1; -; Genomic_DNA.
KW Hypothetical protein; Signal.
FT SIGNAL 1 23 Potential.
SQ SEQUENCE 55 AA; 5602 MW; 0B60FDE7E389FABD CRC64;

Query Match 5.1%; Score 6; DB 2; Length 55;
Best Local Similarity 100.0%; Pred. No. 9.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 6 LAALAA 11

RESULT 906
Q5L3F0_GEOKA PRELIMINARY; PRT; 58 AA.
AC Q5L3F0_GEOKA PRELIMINARY;
DT 01-FEB-2005 (TrEMBLrel. 29, Created)
DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Hypothetical conserved protein.
GN OrderedLocusNames=GK0245;
OS Geobacillus kaustophilus.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
OX NCBI_TaxID=1462;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=HTA426;
```

```
RX PubMed=15576355; DOI=10.1093/nar/gkh970;
RA Takami H., Takaki Y., Chee G.-J., Nishi S., Shimamura S., Suzuki H.,
RA Matsui S., Uchiyama I.;
RT "Thermoadaptation trait revealed by the genome sequence of
RT thermophilic Geobacillus kaustophilus.",
RL Nucleic Acids Res. 32:6292-6303(2004).
DR EMBL; BA000043; BAD74530.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 58 AA; 6514 MW; 76C4BC2DC1D2F178 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 38 LAALAA 43

RESULT 907
Q4T2U0_TETNG PRELIMINARY; PRT; 59 AA.
AC Q4T2U0_TETNG PRELIMINARY;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Chromosome undetermined SCAPI0182, whole genome shotgun sequence.
DE (Fragment).
GN ORFNames=GSTENG00008183001;
OS Tetraodon nigroviridis (Green puffer).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraontoidea; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Da Silva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McRwan P., Bosak S.,
RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Winkler P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC EMBL; CAAE01010182; CAP92792.1; -; Genomic_DNA.
FT NON TER 1 1
SQ SEQUENCE 59 AA; 6306 MW; 42B36AE255E54E60 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 86 VLGILQ 91
Db 47 VLGILQ 52

RESULT 908
```

```
Q9MC16_9CAUD PRELIMINARY; PRT; 60 AA.
ID Q9MC16_9CAUD PRELIMINARY; PRT; 60 AA.
AC Q9MC16;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ORF-60D
GN Name=orf-60D;
OS Bacteriophage P-E1bD.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=120163;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20187497; PubMed=10722621;
DOI=10.1128/JAI.68.4.2205-2214.2000;
RA Sandt C.H., Hill C.W.;
RT "Four different genes responsible for nonimmune immunoglobulin-binding activities within a single strain of Escherichia coli.";
RL Infect. Immun. 68:2205-2214(2000).
DR EMBL; AF151675; AAF63042.1; -; Genomic DNA.
SQ SEQUENCE 60 AA; 6688 MW; F3EECAA267C34F54 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 45 LGGVLA 50

RESULT 909
Q88IN6_PSBPK PRELIMINARY; PRT; 60 AA.
ID Q88IN6_PSBPK PRELIMINARY; PRT; 60 AA.
AC Q88IN6;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PP2963;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22423060; PubMed=12534463;
DOI=10.1046/j.1462-2920.2002.00366.x;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
Moazzez A., Utterback T.R., Rizzo M., Lee K., Kossack D., Moestl D.,
Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuemmler B.,
Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AB016785; AAN68571.1; -; Genomic DNA.
TIGR; PP2963; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 60 AA; 6559 MW; 188BB16ADE29D18F CRC64;

Query Match 5.1%; Score 6; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 EVTTST 13
Db 16 EVTTST 21

RESULT 910
AKH2 LOCM1 STANDARD; PRT; 61 AA.
ID AKH2 LOCM1 STANDARD; PRT; 61 AA.
AC P08379; P10617;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-SEP-1996 (Rel. 34, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Adipokinetic prohormone type 2 precursor (Contains: Adipokinetic hormone II (AKH-II); Adipokinetic hormone precursor-related peptide beta chain (APRP-beta) (6 kDa dimeric peptide B)).
OS Locusta migratoria (migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Corpora cardiaca;
RX MEDLINE=96032738; PubMed=7559443; DOI=10.1074/jbc.270.39.23038;
RA Bogerd J., Kooiman F.P., Pijnenburg M.A.P., Hekking L.H., Gudejans R.C., Vander Horst D.J.;
RT "Molecular cloning of three distinct cDNAs, each encoding a different adipokinetic hormone precursor, of the migratory locust, Locusta migratoria. Differential expression of the distinct adipokinetic hormone precursor genes during flight activity.";
RL J. Biol. Chem. 270:23038-23043(1995).
RN [2]
RP PROTEIN SEQUENCE OF 23-30.
RX MEDLINE=86050918; PubMed=4063072;
RA Siebert K., Morgan P., Mordue W.;
RT "Primary structures of locust adipokinetic hormones II.";
RL Biol. Chem. Hoppe-Seyler 366:723-727(1985).
RN [3]
RP PROTEIN SEQUENCE OF 23-30.
RX MEDLINE=86130555; PubMed=3947348;
RA Gaede G., Goldsworthy G.J., Schaffer M.H., Cook J.C., Rinehart K.L. Jr.;
RT "Sequence analyses of adipokinetic hormones II from corpora cardiaca of Schistocerca nitans, Schistocerca gregaria, and Locusta migratoria by fast atom bombardment mass spectrometry.";
RL Biochem. Biophys. Res. Commun. 134:723-730(1986).
RN [4]
RP PROTEIN SEQUENCE OF 34-61.
RX MEDLINE=89276392; PubMed=2731552;
RA Hietter H., Luu B., Goltzene F., Zachary D., Hoffmann J.A., van Dorsselaer A.;
RT "Isolation and structure of two novel 6-kDa dimeric peptides from the corpora cardiaca of the insect Locusta migratoria. Molecular mass determination by mass spectrometry.";
RL Eur. J. Biochem. 182:77-84(1989).
CC -!- FUNCTION: This hormone, released from cells in the corpora cardiaca after the beginning of flight, causes release of diglycerides from the fat body and then stimulates the flight muscles to use these diglycerides as an energy source.
CC -!- SUBUNIT: Adipokinetic hormone precursor-related peptide (APRP) can form three type of disulfide-bond dimers: p1 (alpha-alpha), p2 (alpha-beta), and p3 (beta-beta).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the AKH/HRTH/RPCH family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC
CC EMBL; X86800; CAA60495.1; -; mRNA.
DR PIR; B58652; AKQL2.
DR InterPro; IPR010475; Adipokinetic_2.
DR InterPro; IPR02047; AKH.
DR Pfam; PF06377; Adipokin_hormo; 1.
DR PROSITE; PS00256; AKH; 1.
```

KW Amidation; Cleavage on pair of basic residues;
 KW Direct protein sequencing; Flight; Neuropeptide;
 KW Pyrrolidone carboxylic acid; Signal.
 FT SIGNAL 1 22
 FT CHAIN 23 61 Adipokinetic prohormone type 2.
 FT PEPTIDE 23 30 Adipokinetic hormone II.
 FT PEPTIDE 34 61 Adipokinetic hormone precursor-related
 FT peptide beta chain.
 FT MOD_RES 23 23 Pyrrolidone carboxylic acid.
 FT MOD_RES 30 30 Tryptophan amide (G-31 provides amide
 group).
 FT DISULFID 59 59 Interchain.
 FT SEQUENCE 61 AA; 6588 MW; 3436B59C3712A046 CRC64;
 Query Match 5.1%; Score 6; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 20 VLAALA 25
 DB 15 VLAALA 20
 RESULT 911
 Q4TU62_TETNG
 ID Q4TU62_TETNG PRELIMINARY; PRT; 61 AA.
 AC Q4TU62;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome undetermined SCAF602, whole genome shotgun sequence.
 DE (Fragment).
 DE ORFNames=GSTNG00004363001;
 GN Tetraodon nigroviridis (Green puffer).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontinae; Tetraodon.
 OX NCBI_TaxID=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 Da Silva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 Anthonard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
 Blomont C., Skalli Z., Cattolico L., Poullain J., De Berardinis V.,
 Craud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 Kallis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
 Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 Winkler P., Lander E.S., Weissbach J., Roest Croliis H.,
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype."
 RL Nature 411:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope; Whitehead
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 DR DR EMBL; CAAB01000602; CAF897070.1; -; Genomic_DNA.
 FT NON TER 1 61
 FT NON TER 61 61
 FT SEQUENCE 61 AA; 6858 MW; A5840CA9E05E557E CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 88 GLLQRA 93

DB 34 GLLQRA 39
 RESULT 912
 Q7UK05_RHOBA
 ID Q7UK05_RHOBA PRELIMINARY; PRT; 62 AA.
 AC Q7UK05;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 DE OrderedLocustNames=RB10939;
 GN Rhodopirellula baltica.
 OS Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Firellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
 RA Schleier H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Firellula sp.
 strain 1.1";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 DR EMBL; BX294152; CAD7076.1; -; Genomic DNA.
 GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR003313; AraC binding.
 DR InterPro; IPR007113; Cupin region.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 62 AA; 6660 MW; 6B698C067A604EEC CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 99 VIEPIV 104
 DB 39 VIEPIV 44
 RESULT 913
 O63372_9ACTI
 ID O63372_9ACTI PRELIMINARY; PRT; 62 AA.
 AC O63372;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser oxyrinchus oxyrinchus.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=40147;
 RN [1]
 RP NUCLEOTIDE SEQUENCE
 RA MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
 RA Birstein V.J., Desalle R.;
 RT "Molecular phylogeny of Acipenserinae."
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF006163; AAC08498.1; -; Genomic_DNA.
 DR SMR; O63372; 1-60.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.

DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 SQ SEQUENCE 62 AA; 7098 MW; 48CDFB2BEEA0B250 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 18 LGGVLA 23
 RESULT 914
 O63373_9ACTI
 ID O63373_9ACTI PRELIMINARY; PRT; 62 AA.
 AC O63373;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser oxyrinchus desotoi.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=40146;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae."
 RL Mol. Phylogenet. Evol. 9:141-155 (1998).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF006164; AAC08499.1; -; Genomic_DNA.
 DR SMR; O63373; 1-60.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 SQ SEQUENCE 62 AA; 7098 MW; 48CDFB2BEEA0B250 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 18 LGGVLA 23
 RESULT 915
 O63375_ACISI
 ID O63375_ACISI PRELIMINARY; PRT; 62 AA.

AC O63375;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser sinensis (Chinese sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=61970;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae."
 RL Mol. Phylogenet. Evol. 9:141-155 (1998).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF006166; AAC08501.1; -; Genomic_DNA.
 DR SMR; O63375; 1-60.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; P:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 SQ SEQUENCE 62 AA; 7024 MW; 1C35429A5ECCB257 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 17 LGGVLA 22
 Db 18 LGGVLA 23
 RESULT 916
 O63386_ACIDA
 ID O63386_ACIDA PRELIMINARY; PRT; 62 AA.
 AC O63386;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser dabryanus (Yangtze sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OX NCBI_TaxID=62061;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=98140330; PubMed=9479703; DOI=10.1006/mpev.1997.0443;
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae."
 RL Mol. Phylogenet. Evol. 9:141-155 (1998).
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF006177; AAC08512.1; -; Genomic_DNA.
 DR SMR; O63386; 1-60.
 DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO: 0016020; C:membrane; IEA.
 DR GO: 0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: 0005739; C:mitochondrion; IEA.
 DR GO: 0046872; F:metal ion binding; IEA.
 DR GO: 0016491; F:oxidoreductase activity; IEA.
 DR GO: 0006118; P:electron transport; IEA.
 DR GO: 0006810; P:transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR Pfam: PF00032; Cytochrom_B_C7_1.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 1
 SQ SEQUENCE 62 AA; 7174 MW; 48D5429A45BCC250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 18 LGGVLA 23

RESULT 917
 O63374 ACIPE
 ID O63374 ACIPE PRELIMINARY; PRT; 62 AA.
 AC O63374;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser persicus (Persian sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OC NCBI_TaxID=61968;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae.";
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF006165; AAC08500.1; -; Genomic_DNA.
 DR SMR; O63374; 1-60.
 DR GO: 0016021; C:integral to membrane; IEA.
 DR GO: 0016020; C:membrane; IEA.
 DR GO: 0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: 0005739; C:mitochondrion; IEA.
 DR GO: 0046872; F:metal ion binding; IEA.
 DR GO: 0006118; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 1
 SQ SEQUENCE 62 AA; 7144 MW; 48D5429A45A0B250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 18 LGGVLA 23

RESULT 919
 O63371 ACINU
 ID O63371 ACINU PRELIMINARY; PRT; 62 AA.
 AC O63371;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser nudiiventris (Fringebarbel sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OC NCBI_TaxID=61967;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae.";
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF006165; AAC08500.1; -; Genomic_DNA.
 DR SMR; O63374; 1-60.
 DR GO: 0016021; C:integral to membrane; IEA.
 DR GO: 0016020; C:membrane; IEA.
 DR GO: 0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: 0005739; C:mitochondrion; IEA.
 DR GO: 0046872; F:metal ion binding; IEA.
 DR GO: 0016491; F:oxidoreductase activity; IEA.
 DR GO: 0006118; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 1
 SQ SEQUENCE 62 AA; 7112 MW; 48D5582BEEA0B250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 18 LGGVLA 23

RESULT 918
 O63370 ACINA
 ID O63370 ACINA PRELIMINARY; PRT; 62 AA.
 AC O63370;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser naccarii (Adriatic sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OC NCBI_TaxID=42330;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae.";
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF006161; AAC08496.1; -; Genomic_DNA.
 DR SMR; O63370; 1-60.
 DR GO: 0016021; C:integral to membrane; IEA.
 DR GO: 0016020; C:membrane; IEA.
 DR GO: 0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: 0005739; C:mitochondrion; IEA.
 DR GO: 0046872; F:metal ion binding; IEA.
 DR GO: 0016491; F:oxidoreductase activity; IEA.
 DR GO: 0006118; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 1
 SQ SEQUENCE 62 AA; 7144 MW; 48D5429A45A0B250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 18 LGGVLA 23

RESULT 919
 O63371 ACINU
 ID O63371 ACINU PRELIMINARY; PRT; 62 AA.
 AC O63371;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser nudiiventris (Fringebarbel sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OC NCBI_TaxID=61967;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae.";
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF006165; AAC08500.1; -; Genomic_DNA.
 DR SMR; O63374; 1-60.
 DR GO: 0016021; C:integral to membrane; IEA.
 DR GO: 0016020; C:membrane; IEA.
 DR GO: 0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: 0005739; C:mitochondrion; IEA.
 DR GO: 0046872; F:metal ion binding; IEA.
 DR GO: 0006118; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 1
 SQ SEQUENCE 62 AA; 7144 MW; 48D5429A45A0B250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 18 LGGVLA 23

RESULT 919
 O63371 ACINU
 ID O63371 ACINU PRELIMINARY; PRT; 62 AA.
 AC O63371;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Acipenser nudiiventris (Fringebarbel sturgeon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
 OC Acipenser.
 OC NCBI_TaxID=61967;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Birstein V.J., DeSalle R.;
 RT "Molecular phylogeny of Acipenserinae.";
 RL Mol. Phylogenet. Evol. 9:141-155(1998).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF006165; AAC08500.1; -; Genomic_DNA.
 DR SMR; O63374; 1-60.
 DR GO: 0016021; C:integral to membrane; IEA.
 DR GO: 0016020; C:membrane; IEA.
 DR GO: 0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: 0005739; C:mitochondrion; IEA.
 DR GO: 0046872; F:metal ion binding; IEA.
 DR GO: 0006118; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 1
 SQ SEQUENCE 62 AA; 7144 MW; 48D5429A45A0B250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||
 Db 18 LGGVLA 23

CC cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF06162; AAC08497.1; -; Genomic_DNA.
 DR SMR; O63371; 1-60.
 DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0016020; C: membrane; IEA.
 DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C: mitochondrion; IEA.
 DR GO; GO:0046872; F: metal ion binding; IEA.
 DR GO; GO:0016491; F: oxidoreductase activity; IEA.
 DR GO; GO:0006118; P: electron transport; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrome_b6_C.
 DR PROSITE; PSS1003; CYTB_CTER; 1.
 DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER
 SQ SEQUENCE 62 AA; 7174 MW; 48D5429AA5BCC250 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 18 LGGVLA 23

RESULT 920
 Q6EQV3 ORYSA
 ID Q6EQV3_ORYSA PRELIMINARY; PRT; 63 AA.
 AC Q6EQV3;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Hypothetical protein OSUNB0075E08.21.
 GN Name=OSUNB0075E08.21;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzoae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005653; BAD28967.1; -; Genomic_DNA.
 DR Gramene; Q6EQV3; -;
 KW Hypothetical protein.
 SQ SEQUENCE 63 AA; 6915 MW; D3C6326C225168A5 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLQRAT 94

Db 36 LLQRAT 41

RESULT 921
 Q4HDV7 CAMCO
 ID Q4HDV7_CAMCO PRELIMINARY; PRT; 63 AA.
 AC Q4HDV7;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Probable periplasmic protein Cj1513c-related protein.
 GN ORFNames=CC01620;
 OS Campylobacter coli RM2228.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=306254;

RNA NUCLEOTIDE SEQUENCE.
 RC STRAIN=RM2228;
 RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
 RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
 RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
 RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
 RA Nelson K.E.;
 RT "Major structural and novel potential virulence mechanisms from the
 RT genomes of multiple Campylobacter species.";
 RL Submitted (DSC-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAFJ01000014; EAL56039.1; -; Genomic DNA.
 SQ SEQUENCE 63 AA; 7239 MW; 593A1250FF204112 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 KEVLVQ 57

Db 45 KEVLVQ 50

RESULT 922
 Q4HR59 CAMUP
 ID Q4HR59_CAMUP PRELIMINARY; PRT; 63 AA.
 AC Q4HR59;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Probable periplasmic protein Cj1513c-related protein.
 GN ORFNames=CUP1307;
 OS Campylobacter upsaliensis RM3195.
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
 OC Campylobacteraceae; Campylobacter.
 OX NCBI_TaxID=306264;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=RM3195;
 RA Fouts D.E., Mongodin E.F., Mandrell R.E., Miller W.G., Rasko D.A.,
 RA Jacques R.J., Brinkac L.M., DeBoy R.T., Parker C.T., Daugherty S.C.,
 RA Dodson R.J., Durkin A.S., Madupu R.R., Sullivan S.A., Shetty J.U.,
 RA Ayodeji M.A., Shvartsbeyn A.A., Schatz M.C., Badger J.H., Fraser C.M.,
 RA Nelson K.E.;
 RT "Major structural and novel potential virulence mechanisms from the
 RT genomes of multiple Campylobacter species.";
 RL Submitted (DSC-2004) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAFJ01000005; EAL53165.1; -; Genomic DNA.
 SQ SEQUENCE 63 AA; 7178 MW; 3A227C869BEA5CA3 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 KEVLVQ 57

Db 45 KEVLVQ 50

RESULT 923
 Q7VTJ1 BORPE
 ID Q7VTJ1_BORPE PRELIMINARY; PRT; 63 AA.
 AC Q7VTJ1;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.

OrderedLocusNames=BP3545;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 ON NCBI_TaxID=520;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Seeger K.,
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640421; CA343804.1; -; Genomic DNA.
 DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
 DR InterPro; IPR005064; UPF0065.
 DR Pfam; PF03401; Bug; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 63 AA; 6883 MW; 27949BBD5078A45F CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 GVLAAL 24
 Db 11 GVLAAL 16
 |||||
 RESULT 924
 Q7W7H0_BORPA
 ID Q7W7H0_BORPA PRELIMINARY; PRT; 63 AA.
 AC Q7W7H0;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BP3549;
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 ON NCBI_TaxID=519;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Seeger K.,
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640430; CA337843.1; -; Genomic DNA.
 DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
 DR InterPro; IPR005064; UPF0065.
 DR Pfam; PF03401; Bug; 1.
 KW Complete proteome; Hypothetical protein.

SQ SEQUENCE 63 AA; 6883 MW; 27949BBD5078A45F CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 GVLAAL 24
 Db 11 GVLAAL 16
 |||||
 RESULT 925
 Q7WKV9_BORBR
 ID Q7WKV9_BORBR PRELIMINARY; PRT; 63 AA.
 AC Q7WKV9;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BB1994;
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 ON NCBI_TaxID=518;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=RB50 / ATCC BAA-588;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
 Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
 Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
 Rabinowitsch E., Rutter S., Sanders M., Saunders R., Squares S., Seeger K.,
 Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,
 Bordetella parapertussis and Bordetella bronchiseptica.";
 RL Nat. Genet. 35:32-40(2003).
 DR EMBL; BX640443; CA32491.1; -; Genomic DNA.
 DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. .; IEA.
 DR InterPro; IPR005064; UPF0065.
 DR Pfam; PF03401; Bug; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 63 AA; 6883 MW; 27949BBD5078A45F CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 63;
 Best Local Similarity 100.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 19 GVLAAL 24
 Db 11 GVLAAL 16
 |||||
 RESULT 926
 Q85DJ6_APOSY
 ID Q85DJ6_APOSY PRELIMINARY; PRT; 63 AA.
 AC Q85DJ6;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Apodemus sylvaticus (European woodmouse).
 OC Mitochondrion.
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Apodemus.
 ON NCBI_TaxID=10129;
 RN [1]

```

RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22563866; PubMed=12675824;
RX DOI=10.1046/j.1365-294X.2003.01752.x;
RA Michaux J.R., Magnanou E., Paradis E., Nieberding C., Libois R.;
RT "Mitochondrial phylogeography of the Woodmouse (Apodemus sylvaticus)
in the Western Palearctic region.";
RL Mol. Ecol. 12:685-697(2003).
CC -!- SUBUNIT: the main subunits of complex b-cl are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AJ511979; CAD54399.1; -; Genomic_DNA.
DR SMR; Q85DJ6; 1-63.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b_C.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 63
SQ SEQUENCE 63 AA; 7201 MW; 3C4EA3288921D96F CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 27 LGGVLA 32

RESULT 927
Q52262_9VIRU
ID Q52262_9VIRU PRELIMINARY; PRT; 63 AA.
AC Q52262;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Triple gene block 3 protein.
GN Name=TGB3;
OS Alternanthera mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxID=85454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ALTMV-SP, and ALTMV-BR;
RA Hammond J., Reinsel M.D., Maroon-Lango C.J.;
RT "Identification of potexvirus isolates from creeping phlox and
trailing portulaca as strains of Alternanthera mosaic virus, and
comparison of the 3-terminal portion of the viral genomes.";
RL Acta Hort. 0:0-0(2005).
DR EMBL; AY850931; AAX86023.1; -; Genomic RNA.
DR EMBL; AY850928; AAX86014.1; -; Genomic RNA.
SQ SEQUENCE 63 AA; 6465 MW; 6CCB28C69AF7BF6C CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 16 GVLAAAL 21

RESULT 928
Q52265_9VIRU
ID Q52265_9VIRU PRELIMINARY; PRT; 63 AA.
AC Q52265;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Triple gene block 3 protein.
GN Name=TGB3;
OS Alternanthera mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxID=85454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ALTMV-SP, and ALTMV-BR;
RA Hammond J., Reinsel M.D., Maroon-Lango C.J.;
RT "Identification of potexvirus isolates from creeping phlox and
trailing portulaca as strains of Alternanthera mosaic virus, and
comparison of the 3-terminal portion of the viral genomes.";
RL Acta Hort. 0:0-0(2005).
DR EMBL; AY850931; AAX86023.1; -; Genomic RNA.
DR EMBL; AY850928; AAX86014.1; -; Genomic RNA.
SQ SEQUENCE 63 AA; 6423 MW; 8CDE28C69AE0782D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 16 GVLAAAL 21

RESULT 929
Q52267_9VIRU
ID Q52267_9VIRU PRELIMINARY; PRT; 63 AA.
AC Q52267;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Triple gene block 3 protein.
GN Name=TGB3;
OS Alternanthera mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flexiviridae;
OC Potexvirus.
OX NCBI_TaxID=85454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=ALTMV-PA;
RA Hammond J., Reinsel M.D., Maroon-Lango C.J.;
RT "Identification of potexvirus isolates from creeping phlox and
trailing portulaca as strains of Alternanthera mosaic virus, and
comparison of the 3-terminal portion of the viral genomes.";
RL Acta Hort. 0:0-0(2005).
DR EMBL; AY850929; AAX86017.1; -; Genomic RNA.
SQ SEQUENCE 63 AA; 6466 MW; 6CCB28C69AFAD29C CRC64;

Query Match 5.1%; Score 6; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAAL 24
Db 16 GVLAAAL 21

RESULT 930
Q4TSF8_9SPHN
ID Q4TSF8_9SPHN PRELIMINARY; PRT; 64 AA.
AC Q4TSF8;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DE Hypothetical protein.
GN ORFNames=ELI0218;

```

OS Erythrobacter litoralis HTCC2594.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
 OC Sphingomonadaceae; Erythrobacter.
 OX NCBI_TaxID=314225;

[1]

RN NUCLEOTIDE SEQUENCE.
 RC STRAIN=HTCC2594;
 RA Giovanni S.J., Cho J.-C., Ferriera S., Johnson J., Kravitz S.,
 RA Halpern A., Remington K., Beeson K., Tran B., Rogers Y.-H.,
 RA Friedman R., Venter J.C.;
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAGG01000001; EAL76412.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 64 AA; 6892 MW; 4B8585B2FC78BA48 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 64;

Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

Db 12 LAALAA 17

RESULT 931

ID Q4JB65 SULAC PRELIMINARY; PRT; 65 AA.
 AC Q4JB65;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocuNames=Saci.0572;
 OS Sulfolobus acidocaldarius.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2285;
 [1]
 RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 33909 / NCIB 11780 / DSM 639;
 RX PubMed=15995215; DOI=10.1126/JB.187.14.4992-4999.2005;
 RA Chen L., Bruegger K., Skovgaard M., Redder P., She Q., Torarinsson E.,
 RA Greve B., Aweyer M., Zibat A., Klenk H.-P., Garrett R.A.;
 RT "The genome of Sulfolobus acidocaldarius, a model organism of the
 RT Crenarchaeota."
 RL J. Bacteriol. 187:4992-4999(2005).
 DR EMBL; CP000077; AAY79964.1; -; Genomic_DNA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 65 AA; 7296 MW; D6DE3BB143DF8F51 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 65;

Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 43 GKGPAI 48

Db 36 GKGPAI 41

RESULT 932

ID Q92N90 RHIME PRELIMINARY; PRT; 65 AA.
 AC Q92N90;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE HYPOTHETICAL TRANSMEMBRANE PROTEIN.
 GN OrderedLocuNames=R02326; ORFNames=SMC01580;
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.

OX NCBI_TaxID=382;

[1]

RN NUCLEOTIDE SEQUENCE.

RC STRAIN=1021;
 RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Goidie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 RA Pohl T., Portetelle D., Puhler A., Purnelle B., Rameperger U.,
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT Sinorhizobium meliloti strain 1021."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 DR EMBL; AL591790; CAC46905.1; -; Genomic_DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 KW Complete proteome; Hypothetical protein; Transmembrane.
 SQ SEQUENCE 65 AA; 7017 MW; 23B0D177FB2BAFDA CRC64;

Query Match 5.1%; Score 6; DB 2; Length 65;

Best Local Similarity 100.0%; Pred. No. 1.1e+03;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26

Db 25 LAALAA 30

RESULT 933

ID Q4S2D5 TETNG PRELIMINARY; PRT; 65 AA.
 AC Q4S2D5;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Chromosome undetermined SCAFI4764, whole genome shotgun sequence.
 DE (Fragment).
 GN ORFNames=GSTENG0025135001;
 OS Tetraodon nigroviridis (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraontoidea; Tetraodontidae; Tetraodon.
 OX NCBI_TaxID=99883;

[1]

RN NUCLEOTIDE SEQUENCE.

RA Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
 RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McGwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
 RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
 RT the early vertebrate proto-karyotype.";
 RL Nature 431:946-957(2004).

[2]

RN NUCLEOTIDE SEQUENCE.

RG Genoscope; Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; CAAB01014764; CAG05197.1; -; Genomic_DNA.

FT NON TER 1

SQ SEQUENCE 65 AA; 7003 MW; 0A8434183F9ASAC1 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 65;

```

Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VVIVGH 39
DB 5 VVIVGH 10

RESULT 934
O05613 PSBP PRELIMINARY; PRT; 67 AA.
AC O05613
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Urf-1 protein.
GN Name=urf-1;
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=KHP41;
RX MEDLINE=97419493; PubMed=9274008;
RA Kholidi G.Y., Yurieva O.V., Gorlenko Z.M., Mindlin S.Z., Bass I.A.,
RA Lomovskaya O.L., Kopteva A.V., Nikiforov V.G.;
RT "Tn5041 : a chimeric mercury resistance transposon closely related to
RT the toluene degradative transposon Tn4651." ;
RL Microbiology 143:2549-2556 (1997).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=KHP41;
RA Kholidi G.Y., Mindlin S.Z., Gorlenko Z.M., Bass I.A., Kalyaeva E.S.,
RA Nikiforov V.;
RT "Host-dependent transposition of Tn5041." ;
RL Russ. J. Genet. 36:365-373 (2000).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=KHP41;
RX MEDLINE=22315381; PubMed=12427948;
RA Kholidi G., Gorlenko Z., Mindlin S., Hobman J., Nikiforov V.;
RT "Tn5041-like transposons: molecular diversity, evolutionary
RT relationships and distribution of distinct variants in environmental
RT bacteria." ;
RL Microbiology 148:3569-3582 (2002).
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=KHP41;
RX PubMed=10822806;
RA Kholidi G., Mindlin S.Z., Gorlenko Z.M., Bass I.A., Kalyaeva E.S.,
RA Nikiforov V.G.;
RT "Molecular genetic analysis of the Tn5041 transposition system." ;
RL Genetika 36:459-469 (2000).
DR EMBL; X98999; CAA67454.1; -; Genomic_DNA.
DR InterPro; IPR007746; MerE.
DR Pfam; PF05052; MerE; 1.
SQ SEQUENCE 67 AA; 7377 MW; 1804BDE8074EB436 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 GVLAAL 24
DB 10 GVLAAL 15

RESULT 935
O7ME50 VIBVY PRELIMINARY; PRT; 67 AA.
AC O7ME50;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)

```

```

DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein VVA0833.
GN OrderedLocusNames=VVA0833;
OS Vibrio vulnificus (strain XJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrrio.
OX NCBI_TaxID=196600;
RN NUCLEOTIDE SEQUENCE.
RP PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen." ;
RL Genome Res. 13:2577-2587 (2003).
DR EMBL; BA000038; BAC96860.1; -; Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 67 AA; 7975 MW; 5BEF6C4DF5EBD4FF CRC64;

Query Match 5.1%; Score 6; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 VLLGGV 20
DB 52 VLLGGV 57

RESULT 936
Q4HCD9_9DEIO PRELIMINARY; PRT; 68 AA.
AC Q4HCD9;
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=DgeODRAFT 2114;
OS Deinococcus geothermalis DSM 11300.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=319795;
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Israni S., Pitluck S., Richardson P.;
RT "Sequencing of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300." ;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 11300;
RG US DOE Joint Genome Institute (JGI-ORNL);
RA Larimer F., Land M.;
RT "Annotation of the draft genome assembly of Deinococcus geothermalis
RT DSM 11300." ;
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHE01000001; EAL84008.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 68 AA; 7693 MW; 7A1068938B79C8F3 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 QOAVIE 101
DB 17 QOAVIE 22

```

```
RESULT 937
Q9Q8X6_9POXV AC Q9Q8X6; PRT; 68 AA.
ID Q9Q8X6_9POXV PRELIMINARY;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Gp102L.
GN Name=s102L;
OS Rabbit fibroma virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Leporipoxvirus.
OX NCBI_TaxID=10271;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Kasza;
RX MEDLINE=84165064; PubMed=6323741;
RA Dalange A.M., Macaulay C., Block W., Mueller T., McFadden G.;
RT "Tumorigenic poxviruses: construction of the composite physical map of
the Shope fibroma virus genome.";
RL J. Virol. 50:408-416(1984).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Kasza;
RX MEDLINE=52074222; PubMed=1660196;
RA Strayer D.S., Jerng H.H., O'Connor K.;
RT "Sequence and analysis of a portion of the genomes of Shope fibroma
virus and malignant rabbit fibroma virus that is important for viral
replication in lymphocytes.";
RL Virology 185:585-595(1991).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Kasza;
RX MEDLINE=20032074; PubMed=10562495; DOI=10.1006/viro.1999.0002;
RA Willer D.O., McFadden G., Evans D.H.;
RT "The complete genome sequence of shope (Rabbit) fibroma virus.";
RL Virology 264:319-343(1999).
DR EMBL; AF170722; AAF17985.1; -; Genomic DNA.
DR InterPro; IPR009236; Chordopox_A13L.
DR Pfam; PF05961; Chordopox_A13L.1.
SQ SEQUENCE 68 AA; 7814 MW; 0532A58336F27E7D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 33 CWIVG 38
Db 11 CWIVG 16

RESULT 938
YJBJ_ECO57 AC P68207; P32691; PRT; 69 AA.
ID YJBJ_ECO57 STANDARD;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE UPF0337 protein yjBj.
GN Name=yjBj; OrderedLocusNames=z5644, ECs5028;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
```

```
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O157:H7 / Sakai / RIMD 0509952 / EHEC;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- SIMILARITY: Belongs to the UPF0337 (csbd) family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC
DR EMBL; AE005174; AAG59244.1; -; Genomic DNA.
DR EMBL; BA000007; BAB38451.1; -; Genomic DNA.
DR PIR; D91257; D91257.
DR PIR; H86097; H86097.
DR SMR; P68207; 1-69.
DR InterPro; IPR008462; CsbD.
DR Pfam; PF05532; CsbD; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 8325 MW; 9A7B50816959F031 CRC64;

Query Match 5.1%; Score 6; DB 1; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QPKGKV 86
Db 12 QPKGKV 17

RESULT 939
YJBJ_ECOL6 AC Q8FB32; PRT; 69 AA.
ID YJBJ_ECOL6 STANDARD;
DT 25-OCT-2004 (Rel. 45, Created)
DT 25-OCT-2004 (Rel. 45, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE UPF0337 protein yjBj.
GN Name=yjBj; OrderedLocusNames=c5016;
OS Escherichia coli O6.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=217992;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;
RX MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Blattner F.R.;
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
CC -1- SIMILARITY: Belongs to the UPF0337 (csbd) family.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
```



```
CC removed.
CC EMBL; AEO16770; AAN83442.1; ALT_INIT; Genomic_DNA.
DR HSSP; P32691; LJYG.
DR SMR; Q8FB32; 1-69.
DR InterPro; IPR008462; CsbD.
DR Pfam; PF05532; CsbD; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 8349 MW; 9A69CD5ADC9E2A1 CRC64;

Query Match
Best Local Similarity 5.1%; Score 6; DB 1; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QPKGKV 86
DB 12 QPKGKV 17

RESULT 940
YJBJ ECOLI
ID YJBJ ECOLI STANDARD; PRT; 69 AA.
AC P68206; P32691;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE UPF0337 protein yjbj.
GN Name=yjbj; OrderedLocusNames=b4045;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
RN [2]
RP PROTEIN SEQUENCE OF 1-12.
RC STRAIN=K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12."
RL Electrophoresis 18:1259-1313(1997).
RN [3]
RP PROTEIN SEQUENCE OF 1-10.
RC STRAIN=K12;
RX MEDLINE=93085675; PubMed=9868784;
RA Wainger V.C., Humphrey-Smith I.;
RT "Small genes/gene-products in Escherichia coli K-12."
RL FEMS Microbiol. Lett. 169:375-382(1998).
RN [4]
RP STRUCTURE BY NMR.
RX PubMed=12001338; DOI=10.1002/prot.10120;
RA Pineda-Lucena A., Liao J., Wu B., Yee A., Cort J.R., Kennedy M.A.,
RA Edwards A.M., Arrowsmith C.H.;
RT "NMR structure of the hypothetical protein encoded by the yjbj gene
RT from Escherichia coli."
RL Proteins 47:572-574(2002).
RN [5]
RP INTERACTION:
CC P08622:dnaJ; NbExp=1; Interact=EBI-549456, EBI-545285;
CC -1- SIMILARITY: Belongs to the UPF0337 (csbD) family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC as long as its content is in no way modified and this statement is not
CC removed.
```

```
CC EMBL; U00006; AAC43139.1; -; Unassigned_DNA.
DR EMBL; U00096; AAC77015.1; -; Genomic_DNA.
DR PIR; D65212; D65212.
DR PDB; 1RYK; NMR; A=1-69.
DR SMR; P68206; 1-69.
DR InterAct; P68206; -.
DR EcoBASE; EBI872; -.
DR EcoGene; EGI1928; yjbj.
DR InterPro; IPR008462; CsbD.
DR Pfam; PF05532; CsbD; 1.
KW 3D-structure; Complete proteome; Direct protein sequencing.
FT TURN 6
FT TURN 9
FT TURN 10
FT TURN 21
FT TURN 26
FT TURN 33
FT TURN 34
FT TURN 36
FT TURN 46
FT TURN 51
FT TURN 65
SQ SEQUENCE 69 AA; 8325 MW; 9A7B50816959F031 CRC64;

Query Match
Best Local Similarity 5.1%; Score 6; DB 1; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QPKGKV 86
DB 12 QPKGKV 17

RESULT 941
YJBJ SHIFL
ID YJBJ SHIFL STANDARD; PRT; 69 AA.
AC P68208; P32691;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DE UPF0337 protein yjbj.
GN Name=yjbj; OrderedLocusNames=SF4160, S3571;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157."
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA DOI=10.1128/JAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T."
RL Infect. Immun. 71:2775-2786(2003).
RN [3]
RP -1- SIMILARITY: Belongs to the UPF0337 (csbD) family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; A005674; AAN45582.1; ALT_INIT; Genomic_DNA.
DR EMBL; A016989; AAP18617.1; -; Genomic_DNA.
DR SNR; P68208; 1-69.
DR InterPro; IPR008462; Cebd.
DR Pfam; PF05532; Cebd; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 8325 MW; 9A7B50816959F031 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKV 86
Db 12 QFKGKV 17

RESULT 942
Q9CP44_PASMU PRELIMINARY; PRT; 69 AA.
AC Q9CP44;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DB MopI.
GN Names:mopI; OrderedLocusNames:PM0216;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida Pm70.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; A006055; AAK02300.1; -; Genomic_DNA.
DR HSP; P08854; 1GVO.
DR GO; GO:0030151; P:molybdenum ion binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004606; Mop.
DR InterPro; IPR005116; TOBE.
DR Pfam; PF03459; TOBE; 1.
DR TIGRFAMs; TIGR00638; Mop; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 7135 MW; E770844488A57DFA CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKV 86
Db 8 QFKGKV 13

RESULT 943
Q99570_ASTDO PRELIMINARY; PRT; 69 AA.
AC Q99570;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Athenes dorbignyi (Creamy-breasted canastero).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae; Asthenes.
OX NCBI_TaxID=9138;

CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; A005674; AAN45582.1; ALT_INIT; Genomic_DNA.
DR EMBL; A016989; AAP18617.1; -; Genomic_DNA.
DR SNR; P68208; 1-69.
DR InterPro; IPR008462; Cebd.
DR Pfam; PF05532; Cebd; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 8325 MW; 9A7B50816959F031 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKV 86
Db 12 QFKGKV 17

RESULT 942
Q9CP44_PASMU PRELIMINARY; PRT; 69 AA.
AC Q9CP44;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DB MopI.
GN Names:mopI; OrderedLocusNames:PM0216;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida Pm70.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; A006055; AAK02300.1; -; Genomic_DNA.
DR HSP; P08854; 1GVO.
DR GO; GO:0030151; P:molybdenum ion binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004606; Mop.
DR InterPro; IPR005116; TOBE.
DR Pfam; PF03459; TOBE; 1.
DR TIGRFAMs; TIGR00638; Mop; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 7135 MW; E770844488A57DFA CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKV 86
Db 8 QFKGKV 13

RESULT 943
Q99570_ASTDO PRELIMINARY; PRT; 69 AA.
AC Q99570;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Athenes dorbignyi (Creamy-breasted canastero).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae; Asthenes.
OX NCBI_TaxID=9138;

CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; A005674; AAN45582.1; ALT_INIT; Genomic_DNA.
DR EMBL; A016989; AAP18617.1; -; Genomic_DNA.
DR SNR; P68208; 1-69.
DR InterPro; IPR008462; Cebd.
DR Pfam; PF05532; Cebd; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 8325 MW; 9A7B50816959F031 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 1; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKV 86
Db 12 QFKGKV 17

RESULT 942
Q9CP44_PASMU PRELIMINARY; PRT; 69 AA.
AC Q9CP44;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DB MopI.
GN Names:mopI; OrderedLocusNames:PM0216;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida Pm70.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; A006055; AAK02300.1; -; Genomic_DNA.
DR HSP; P08854; 1GVO.
DR GO; GO:0030151; P:molybdenum ion binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR004606; Mop.
DR InterPro; IPR005116; TOBE.
DR Pfam; PF03459; TOBE; 1.
DR TIGRFAMs; TIGR00638; Mop; 1.
KW Complete proteome.
SQ SEQUENCE 69 AA; 7135 MW; E770844488A57DFA CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 81 QFKGKV 86
Db 8 QFKGKV 13

RESULT 943
Q99570_ASTDO PRELIMINARY; PRT; 69 AA.
AC Q99570;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Athenes dorbignyi (Creamy-breasted canastero).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae; Asthenes.
OX NCBI_TaxID=9138;
```

```
RN NUCLEOTIDE SEQUENCE.
RP MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053803; AAD09972.1; -; Genomic_DNA.
DR SMR; O99570; 2-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
SQ SEQUENCE 69 AA; 7861 MW; 6A8B3ED38D020400 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22

RESULT 944
Q99560_9FURN PRELIMINARY; PRT; 69 AA.
AC Q99560;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca demissa.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=86270;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053793; AAD09962.1; -; Genomic_DNA.
DR SMR; O99560; 1-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS1003; CYTB_CTER; 1.
```

KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1
 FT NON TER 69
 SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||

Db 17 LGGVLA 22

RESULT 945

OS9563 9FURN
 ID OS9563 9FURN PRELIMINARY; PRT; 69 AA.
 AC OS9563
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Cranioleuca pyrrhophia.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86275;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 RT relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF053796; AAD09965.1; -; Genomic_DNA.
 DR EMBL; AF053796; AAD09965.1; -; Genomic_DNA.
 DR SMR; O99563; 1-69.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B_C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1
 FT NON TER 69
 SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||

Db 17 LGGVLA 22

RESULT 946

OS9559 9FURN
 ID OS9559 9FURN PRELIMINARY; PRT; 69 AA.
 AC OS9559
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)

DE Cytochrome b (Fragment).
 OS Cranioleuca erythropis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86271;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 RT relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF053792; AAD09961.1; -; Genomic_DNA.
 DR EMBL; AF053792; AAD09961.1; -; Genomic_DNA.
 DR SMR; O99559; 1-69.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B_C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1
 FT NON TER 69
 SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 |||||

Db 17 LGGVLA 22

RESULT 947

OS9568 9FURN
 ID OS9568 9FURN PRELIMINARY; PRT; 69 AA.
 AC OS9568
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Cranioleuca vulpina vulpula.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86277;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 RT relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 CC EMBL; AF053801; AAD09970.1; -; Genomic_DNA.
 DR EMBL; AF053801; AAD09970.1; -; Genomic_DNA.
 DR SMR; O99568; 1-69.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.

DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006118; P:electron transport; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR Pfam: PF00032; Cytochrom_B_C; 1.
 DR PROSITE: PS1003; Cytb_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 69 69
 SQ SEQUENCE 69 AA; 7940 MW; BD8BEC40849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 17 LGGVLA 22

RESULT 948
 O99558_9FURN PRELIMINARY; PRT; 69 AA.
 AC O99558
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Cranioleuca antisiensis.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archoosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86267;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL: AF053791; AAD09960.1; -; Genomic_DNA.
 DR SMR; O99558; 1-69.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0006118; P:oxidoreductase activity; IEA.
 DR GO: GO:0006810; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; Cytb_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 69 69
 SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 17 LGGVLA 22

RESULT 949
 O99564_9FURN PRELIMINARY; PRT; 69 AA.
 AC O99564
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Cranioleuca obsoleta.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archoosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86274;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL: AF053797; AAD09966.1; -; Genomic_DNA.
 DR SMR; O99564; 1-69.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0006118; P:oxidoreductase activity; IEA.
 DR GO: GO:0006810; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; Cytb_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 69 69
 SQ SEQUENCE 69 AA; 7940 MW; BD8BEC40849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 17 LGGVLA 22

RESULT 949
 O99564_9FURN PRELIMINARY; PRT; 69 AA.
 AC O99564
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Cranioleuca obsoleta.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archoosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86274;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL: AF053797; AAD09966.1; -; Genomic_DNA.
 DR SMR; O99564; 1-69.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0006118; P:oxidoreductase activity; IEA.
 DR GO: GO:0006810; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; Cytb_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 69 69
 SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 17 LGGVLA 22

RESULT 950
 O99562_9FURN PRELIMINARY; PRT; 69 AA.
 AC O99562
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 OS Cranioleuca pyrrhophia.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archoosauria; Aves; Neognathae; Passeriformes; Furnariidae;
 OC Cranioleuca.
 OX NCBI_TaxID=86275;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
 RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
 RT "A case of rapid diversification in the neotropics: phylogenetic
 relationships among Cranioleuca spinetails (Aves, Furnariidae).";
 RL Mol. Phylogenet. Evol. 12:273-281(1999).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL: AF053791; AAD09960.1; -; Genomic_DNA.
 DR SMR; O99558; 1-69.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO: GO:0005739; C:mitochondrion; IEA.
 DR GO: GO:0046872; F:metal ion binding; IEA.
 DR GO: GO:0006118; P:oxidoreductase activity; IEA.
 DR GO: GO:0006810; P:electron transport; IEA.
 DR InterPro: IPR005798; Cytb_b6_C.
 DR PROSITE: PS1003; Cytb_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1 1
 FT NON_TER 69 69
 SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
 |||||
 Db 17 LGGVLA 22

```

CC EMBL; AF053795; AAD09964.1; -; Genomic_DNA.
DR SMR; 099562; 1-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
SQ SEQUENCE 69 AA; 7940 MW; BD8BEC40849B99409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22
|||||

RESULT 951
ID O99571_9DEND PRELIMINARY; PRT; 69 AA.
AC O99571;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Xiphorhynchus fuscus.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Dendrocolaptidae;
OC Xiphorhynchus.
OX NCBI_TaxID=326918;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053798; AAD09967.1; -; Genomic_DNA.
DR SMR; 099565; 1-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
SQ SEQUENCE 69 AA; 7954 MW; AA34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22
|||||

RESULT 953
ID O99584_9FURN PRELIMINARY; PRT; 69 AA.
AC O99584;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca curtata.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=85407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
RT "A case of rapid diversification in the Neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053804; AAD09973.1; -; Genomic_DNA.
DR SMR; 099571; 1-69.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005739; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
SQ SEQUENCE 69 AA; 7885 MW; 6A482BBE68603A2E3 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22
|||||

```

```

Db 17 LGGVLA 22

RESULT 952
ID O99565_CRAHE PRELIMINARY; PRT; 69 AA.
AC O99565;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca henricae (Bolivian spinetail).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=86272;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053798; AAD09967.1; -; Genomic_DNA.
DR SMR; 099565; 1-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
SQ SEQUENCE 69 AA; 7954 MW; AA34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22
|||||

RESULT 953
ID O99584_9FURN PRELIMINARY; PRT; 69 AA.
AC O99584;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca curtata.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=85407;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
RT "A case of rapid diversification in the Neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053804; AAD09973.1; -; Genomic_DNA.
DR SMR; 099571; 1-69.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0005739; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C; 1.
DR PROSITE; PS51003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
SQ SEQUENCE 69 AA; 7885 MW; 6A482BBE68603A2E3 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
Db 17 LGGVLA 22
|||||

```

RL Mol. Phylogenet. Evol. 0:0-0(1999).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053225; AAD04820.1; -; Genomic_DNA.
DR SMR; O95584; 1-69.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrion; IEA.
DR GO; GO:0046872; F: metal ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C1.
DR PROSITE; PS1003; CYTB_CTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 69 69
SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 17 LGGVLA 22

RESULT 954
O99566 9FURN
ID O99566 9FURN PRELIMINARY; PRT; 69 AA.
AC O99566;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca marcapatae.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=86273;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
RT relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053799; AAD09968.1; -; Genomic_DNA.
DR SMR; O99566; 1-69.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrion; IEA.
DR GO; GO:0046872; F: metal ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR PROSITE; PS1003; CYTB_CTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 69 69
SQ SEQUENCE 69 AA; 7937 MW; F4B558D3849B9E62 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 17 LGGVLA 22
Db 17 LGGVLA 22
RESULT 955
O99567 9FURN
ID O99567 9FURN PRELIMINARY; PRT; 69 AA.
AC O99567;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca albiceps.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=86268;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arcander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
RT relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053800; AAD09969.1; -; Genomic_DNA.
DR SMR; O99567; 1-69.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005746; C: mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C: mitochondrion; IEA.
DR GO; GO:0046872; F: metal ion binding; IEA.
DR GO; GO:0016491; F: oxidoreductase activity; IEA.
DR GO; GO:0006118; P: electron transport; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom_B_C1.
DR PROSITE; PS1003; CYTB_CTER; 1.
DR Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1 1
FT NON_TER 69 69
SQ SEQUENCE 69 AA; 7984 MW; A4A34BD6849E9109 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22

Db 17 LGGVLA 22

RESULT 956
O99561 9FURN
ID O99561 9FURN PRELIMINARY; PRT; 69 AA.
AC O99561;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca albicapilla.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.

```

OX NCBI_TaxID=86266;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053794; AAD09958.1; -; Genomic_DNA.
DR SMR; O99561; 1-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0046020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.
DR Pfam; PF00032; Cytochrom B_C; 1.
DR PROSITE; PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 69
FT NON_TER 69
SQ SEQUENCE 69 AA; 7954 MW; A4A34ED3849B9409 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db |||||
17 LGGVLA 22

RESULT 957
ID Q9ZXP2_9FURN PRELIMINARY; PRT; 69 AA.
AC Q9ZXP2;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-FEB-2005 (TRENBLrel. 29, Last annotation update)
DE Cytochrome b (Fragment).
OS Cranioleuca baroni.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Ruteleostomi;
OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae;
OC Cranioleuca.
OX NCBI_TaxID=86269;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=99343614; PubMed=10413622; DOI=10.1006/mpev.1999.0617;
RA Garcia-Moreno J., Arctander P., Fjeldsa J.;
RT "A case of rapid diversification in the neotropics: phylogenetic
relationships among Cranioleuca spinetails (Aves, Furnariidae).";
RL Mol. Phylogenet. Evol. 12:273-281(1999).
CC -!- SUBUNIT: The main subunits of complex b-cl are: cytochrome b,
cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF053790; AAD09959.1; -; Genomic_DNA.
DR SMR; O99561; 1-69.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0046020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:metal ion binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR005798; Cytb_b6_C.

Query Match 5.1%; Score 6; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db |||||
17 LGGVLA 22

RESULT 958
ID Q74K20_LACJO PRELIMINARY; PRT; 70 AA.
AC Q74K20;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE ATP synthase C chain.
GN OrderedLocusNames=LJ0935;
OS Lactobacillus johnsonii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=33959;
RN NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NCC 533;
RX PubMed=14983040; DOI=10.1073/pnas.0307327101;
RA Pridmore R.D., Berger B., Desiere F., Vilanova D., Barretto C.,
RA Pittet A.-C., Zwanen M.-C., Rouvet M., Altermann E., Barrangou R.,
RA Mollet B., Mercenier A., Klaenhammer T., Arigoni F., Schell M.A.;
RT "The genome sequence of the probiotic intestinal bacterium
Lactobacillus johnsonii NCC 533."
RL Proc. Natl. Acad. Sci. U.S.A. 101:2512-2517(2004).
DR EMBL; AE017202; AAS08756.1; -; Genomic_DNA.
DR HSP; P00844; I491.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0045263; C:proton-transporting ATP synthase complex, c. ; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. ; IEA.
DR GO; GO:0016820; F:hydrolase activity, acting on acid anhydrid. ; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0015986; P:ATP synthesis coupled proton transport; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002379; ATPase_Csub.
DR InterPro; IPR005953; ATP synth C.
DR InterPro; IPR000454; Sub_ATPase_Csub.
DR Pfam; PF00137; ATP-synt C; 1.
DR PRINTS; PR00124; ATPASEC.
DR TIGRFAMs; TIGR01260; ATP synt c; 1.
DR PROSITE; PS00605; ATPASE_C; 1.
KW CPO; Complete proteome; Hydrogen ion transport; Ion transport;
Lipid-binding; Transmembrane; Transport.
SQ SEQUENCE 70 AA; 7220 MW; 141F00874620DF87 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LAALAA 26
Db |||||
12 LAALAA 17

RESULT 959
```

Q9HND9 HALSA

ID Q9HND9 HALSA PRELIMINARY; PRT; 71 AA.
 AC Q9HND9; 5.1%; Score 6; DB 2; Length 71;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE H+-transporting ATP synthase subunit K.
 GN Name=atpK; OrderedLocusNames=VNG2143g;
 OS Halobacterium salinarum (Halobacterium halobium).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=2242;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
 RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
 RA Ng W.V., Kennedy S.P., Mahaias G.G., Berquist B., Pan M.,
 RA Shukla H.D., Leaky S.R., Baliga N.S., Thorsen V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welter R., Goo Y.A.,
 RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonowski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Ikenbarger T.A., Peck R.F., Pohlischroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL; AE005102; AGO2081.1; -; Genomic_DNA.
 DR PIR; E84364; E84364.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
 DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
 DR GO; GO:0046961; F:hydrogen-transporting ATPase activity; IEA.
 DR GO; GO:0015986; F:ATP synthetase coupled proton transport; IEA.
 DR GO; GO:0006811; P:ion transport; IEA.
 DR GO; GO:0015992; P:proton transport; IEA.
 DR InterPro; IPR002379; ATPase Csub.
 DR InterPro; IPR000245; Vac.ATPsynth_Csub.
 DR Pfam; PF00137; ATP-synt_C; 1.
 DR PRINTS; PR00122; VACATPASE.
 DR KW Complete proteome; Hydrogen ion transport; Ion transport;
 KW Transmembrane; Transport.
 SQ SEQUENCE 71 AA; 6957 MW; CA0B15900B14FD86 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 71;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 LALAAA 26
 DB 16 LALAAA 21

RESULT 960

ID Q4Z987_9CAUD PRELIMINARY; PRT; 71 AA.
 AC Q4Z987;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Staphylococcus phage Twort.
 OS Staphylococcus phage Twort.
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
 OX NCBI_TaxID=55510;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC PubMed=15788529; DOI=10.1073/pnas.0501140102;
 RA Kwan T., Liu J., Dubow M., Gros P., Pelletier J.;
 RT "The complete genomes and proteomes of 27 Staphylococcus aureus
 bacteriophages";
 RL Proc. Natl. Acad. Sci. U.S.A. 102:5174-5179(2005).
 DR EMBL; AY954970; AAX92448.1; -; Genomic_DNA.
 KW Transmembrane.
 SQ SEQUENCE 71 AA; 7770 MW; CF4E35DD52113878 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 71;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 EVTTST 13
 DB 20 EVTTST 25

RESULT 961

ID Q4LRG1_9BURK PRELIMINARY; PRT; 71 AA.
 AC Q4LRG1;
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DE Molybdenum-pterin binding protein.
 GN ORFNames=Bcen2424DRAFT_2923;
 OS Burkholderia cenocepacia H12424.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Burkholderia; Burkholderia cenocepacia complex.
 OX NCBI_TaxID=331272;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H12424;
 RA US DOE Joint Genome Institute (JGI-PGF);
 RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
 RA Hammon N., Istrani S., Pitluck S., Richardson P.;
 RT "Sequencing of the draft genome assembly of Burkholderia cenocepacia
 H12424";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H12424;
 RA US DOE Joint Genome Institute (JGI-ORNL);
 RA Larimer F., Land M.;
 RT "Annotation of the draft genome assembly of Burkholderia cenocepacia
 H12424";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AAKL01000024; EAM18594.1; -; Genomic_DNA.
 SQ SEQUENCE 71 AA; 7658 MW; 346F709ED1791167 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 71;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 QFKGKV 86
 DB 11 QFKGKV 16

RESULT 962

ID Q6YQSS_ONYPE PRELIMINARY; PRT; 71 AA.
 AC Q6YQSS;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Uncharacterized BCR.
 GN OrderedLocusNames=PAM298;
 OS Onion yellow phytoplasma.
 OC Bacteria; Firmicutes; Mollicutes; Acholeplasmatales;
 OC Acholeplasmataceae; Candidatus Phytoplasma;
 OC Candidatus Phytoplasma asteris.
 OX NCBI_TaxID=100379;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=OY-M;
 RX PubMed=14661021; DOI=10.1038/ng1277;

RA Oshima K., Kakizawa S., Nishigawa H., Jung H.-Y., Wei W., Suzuki S.,
RA Arashida R., Nakata D., Miyata S.-I., Ugaki M., Namba S.;
RT "Reductive evolution suggested from the complete genome sequence of a
RL plant-pathogenic phytoplasma.";
DR EMBL; AF006628; BAD04383.1; -; Genomic_DNA.
DR InterPro: IPR005359; UPF0154.
DR Pfam: PF03672; UPF0154; 1.
DR ProDom: PD048972; UPF0154; 1.
KW Complete proteome.
SQ SEQUENCE 71 AA; 8405 MW; A8D52B614E238B55 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LGGVLA 21
DB 15 LGGVLA 20
|||||

RESULT 963
O99702_9AVES
ID O99702_9AVES PRELIMINARY; PRT; 73 AA.
AC O99702;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Surniculus lugubris.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Cuculiformes; Cuculidae; Surniculus.
OX NCBI_TaxID=78210;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Blood;
RA Aragon S., Moller A.P., Soler J.J., Soler M.;
RT "Molecular phylogeny of cuckoos supports a polyphyletic origin of
RT brood parasitism.";
RL J. Evol. Biol. 12:495-506(1999).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AF072619; AAC70981.1; -; Genomic_DNA.
DR SMR; O99702; 1-73.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro: IPR005798; Cytb_b_C.
DR Pfam: PF00032; Cytochrom_B_C; 1.
DR PROSITE: PS1003; CYTB_CTER; 1.
KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
PT NON_TER 1
FT NON_TER 73
SQ SEQUENCE 73 AA; 8173 MW; E65726FB9D05E2E CRC64;

Query Match 5.1%; Score 6; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
DB 9 LGGVLA 14
|||||

RESULT 964
Q8MVK3_9ASCI
ID Q8MVK3_9ASCI PRELIMINARY; PRT; 74 AA.
AC Q8MVK3;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Kunitz serine protease inhibitor-like protein (Fragment).
GN Name=kdp;
OS Boltenia villosa.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC Stolidobranchia; Pyruidae; Boltenia.
OX NCBI_TaxID=63515;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=22248966; PubMed=12361966;
RA Davidson B.J., Swalla B.J.;
RT "A molecular analysis of ascidian metamorphosis reveals activation of
RT an innate immune response";
RL Development 129:4739-4751(2002).
CC -!- SIMILARITY: Contains 1 BPTI/Kunitz inhibitor domain.
DR EMBL; AF483058; AAM76138.1; -; mRNA.
DR HSP; P31713; 1SHP.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004867; F:serine-type endopeptidase inhibitor activity; IEA.
DR InterPro: IPR002223; Prot_Inh_Kunz-m.
DR Pfam: PF00014; Kunitz_BPTI; 1.
DR SMART; SM00131; KU; 1.
DR PROSITE; PS50279; BPTI_KUNITZ_2; 1.
KW Protease.
FT NON_TER 1
SQ SEQUENCE 74 AA; 8555 MW; A882A12254DD6F2C CRC64;

Query Match 5.1%; Score 6; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 84 GKVLGL 89
DB 11 GKVLGL 16
|||||

RESULT 965
Q9B6S0_9AVES
ID Q9B6S0_9AVES PRELIMINARY; PRT; 74 AA.
AC Q9B6S0;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Cytochrome b (Fragment).
OS Mullerornis agilis.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Palaeognathae; Struthioniformes; Aepyornithidae;
OC Mullerornis.
OX NCBI_TaxID=147494;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Sub-fossil tibiotarsus;
RX MEDLINE=21085666; PubMed=11217857; DOI=10.1038/35055536;
RA Cooper A., Lalueza-Fox C., Anderson S., Rambaut A., Austin J.,
RA Ward R.;
RT "Complete mitochondrial genome sequences of two extinct moas clarify
RT ratite evolution.";
RL Nature 409:704-707(2001).
CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL; AY016019; AAK08532.1; -; Genomic_DNA.
DR SMR; Q9B6S0; 1-74.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO; GO:0005739; C:mitochondrion; IEA.
DR GO; GO:0046872; F:oxidoreductase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.

OS Gibberella zeae PH-1.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocorycetidae; Hypocreales; Nectriaceae; Gibberella.
 OX NCBI_TaxID=229533;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PH-1;
 RA Birren B., Nusbbaum C., Abouelheil A., Allen N., Anderson S.,
 RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
 RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
 RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., Dearellano K.,
 RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
 RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,
 RA Gardyna S., Gierre S., Graham L., Grand-Pierre N., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
 RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
 RA Kellis C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
 RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
 RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
 RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
 RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
 RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
 RA Rachupka A., Ramaamy U., Raymond C., Retta R., Rise C., Rogov P.,
 RA Roman J., Schauer S., Schupbach R., Seaman S., Severy P., Smirnov S.,
 RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
 RA Talamas J., Teefaye S., Theodore J., Topham K., Travers M.,
 RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Willson B.,
 RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
 RA Lander E.;
 RT "Fusarium graminearum genome sequence."
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AACM01000322; EAA76578.1; -; Genomic_DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 76 AA; 8638 MW; 2C0200B7C9557BDE CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 76;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 90 LQRAQ 95
 DB 18 LQRAQ 23
 RESULT 970
 ID Q75E13 ASHGO PRELIMINARY; PRT; 77 AA.
 AC Q75E13;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE AAR096Wp.
 GN Name=AAR096W;
 OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
 OX NCBI_TaxID=33169;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=ATCC 10895;
 RX PubMed=15001715; DOI=10.1126/science.1095781;
 RA Dietrich F.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
 RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
 RA Gaffney T.D., Philippsen P.;
 RT "The Ashbya gossypii genome as a tool for mapping the ancient
 RT Saccharomycetes cerevisiae genome."
 RL Science 304:304-307(2004).
 DR EMBL; AE016814; AAS0461.1; -; Genomic_DNA.
 DR AGD; AAR096W; -.
 KW Complete proteome.

SQ SEQUENCE 77 AA; 8353 MW; 42B6B952501B2F71 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 77;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAA 26
 DB 44 LAALAA 49
 RESULT 971
 ID Q9RCD9_XANCA PRELIMINARY; PRT; 77 AA.
 AC Q9RCD9;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Urf-1 protein.
 GN Name=urf-1;
 OS Xanthomonas campestris.
 OG Plasmid pKLUH443.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=339;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=TAP44-3;
 RX MEDLINE=99406912; PubMed=10476039;
 RA Minakhina S., Kholodii G., Mindlin S., Yurieva O., Nikiforov V.;
 RT "Tn5053 family transposons are site hunters sensing plasmid res
 RT sites occupied by cognate resolvases."
 RL Mol. Microbiol. 33:1059-1068(1999).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=TAP44-3;
 RX MEDLINE=20331541; PubMed=10875286; DOI=10.1016/S0923-2508(00)00149-2;
 RA Kholodii G., Yurieva O., Mindlin S., Gorlenko Z., Rybochkin V.,
 RA Nikiforov V.;
 RT "Tn5044, a novel Tn3 family transposon coding for temperature
 RT sensitive mercury resistance."
 RL Res. Microbiol. 151:291-302(2000).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=TAP44-3;
 RX MEDLINE=22290346; PubMed=12403178; DOI=10.1023/A:1020185206563;
 RA Kholodii G., Bogdanova E.;
 RT "Tn5044-conferred mercury resistance = depends on temperature: the
 RT complexity of the character of thermosensitivity."
 RL Genetica 115:233-241(2002).
 DR EMBL; Y17691; CAB65708.1; -; Genomic_DNA.
 DR InterPro: IPR007746; MerE.
 DR Pfam: PF05052; MerB; 1.
 KW Plasmid.
 SQ SEQUENCE 77 AA; 8078 MW; E598A410B827093B CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 77;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 19 GVLAA 24
 DB 19 GVLAA 24
 RESULT 972
 ID Q9RGY6_LACAC PRELIMINARY; PRT; 77 AA.
 AC Q9RGY6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, Last annotation update)
 DE F1FO-ATPase subunit c (H+-transporting ATP synthase chain c)

(EC 3.6.3.14).
DE Name=atpE; OrderedLocusNames=LBA0773;
GN Lactobacillus acidophilus
OS Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1579;
RN NCBI_TaxID=1579;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NC56;
RX MEDLINE=99440166; PubMed=10510230;
RA Kullen M.J., Kleenhammer T.R.;
RT "Identification of the pH-inducible, proton-translocating F1F0-ATPase
RT (atpEFHAGDC) operon of Lactobacillus acidophilus by differential
RT display: gene structure, cloning and characterization.";
RL Mol. Microbiol. 33:1152-1161(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=NC56;
RX PubMed=15671160; DOI=10.1073/pnas.0409188102;
RA Altermann E., Russell W.M., Azcarate-Peril M.A., Barrangou R.,
RA Buck B.L., McAniff O., Southern N., Dobson A., Duong T., Callanan M.,
RA Lick S., Hamrick A., Cano R., Kleenhammer T.R.;
RT "Complete genome sequence of the probiotic lactic acid bacterium
RT Lactobacillus acidophilus NCFM.";
RL Proc. Natl. Acad. Sci. U.S.A. 102:3906-3912(2005).
DR EMBL; AF098522; AAF22493.1; -; Genomic DNA.
DR EMBL; CP000033; AAV42639.1; -; Genomic DNA.
DR HSP; P00844; IAA1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016469; C:proton-transporting two-sector ATPase complex; IEA.
DR GO; GO:0046933; F:hydrogen-transporting ATP synthase activity; IEA.
DR GO; GO:0046961; F:hydrogen-transporting ATPase activity, rota. .; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0016820; F:hydrolyase activity, acting on acid anhydrid. .; IEA.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0015986; P:ATP synthetis coupled proton transport; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0015992; P:proton transport; IEA.
DR InterPro; IPR002379; ATPase_Csub.
DR InterPro; IPR005953; ATP_synth_C.
DR InterPro; IPR000454; Eub_ATPase_Csub.
DR Pfam; PF00137; ATP-synt C; 1.
DR PRINTS; PR00124; ATPASEC.
DR TIGRFAMS; TIGR01260; ATP synt_c; 1.
DR PROSITE; PS00605; ATPASE_C; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 77 AA; 7926 MW; EF07386BE4A23336 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.3e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 19 LAALAA 24

RESULT 973
Q92K74 RHIME
ID Q92K74 RHIME PRELIMINARY; PRT; 77 AA.
AC Q92K74;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein SMC04313.
GN OrderedLocusNames=RO1968; ORFNames=SMC04313;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1021;

MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler P., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Rameperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591789; CAC46547.1; -; Genomic DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 77 AA; 8471 MW; 9FC6D733170F41C3 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.3e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
Db 65 LAALAA 70

RESULT 974
Q89YCI BRAJA
ID Q89YCI BRAJA PRELIMINARY; PRT; 77 AA.
AC Q89YCI;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bsr0033 protein.
GN OrderedLocusNames=ber0033;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=USDA 110;
RX MEDLINE=2248498; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110";
RL DNA Res. 9:189-197(2002).
DR EMBL; BA000040; BAC45298.1; -; Genomic DNA.
DR GO; GO:0016810; F:hydrolyase activity, acting on carbon-nitrog. .; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR002509; Polysac_deacet.
DR Pfam; PF01522; Polysacc_deac_1; 1.
KW Complete proteome.
SQ SEQUENCE 77 AA; 8587 MW; C970A6C191E96509 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 1.3e+03; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
Db 32 VLAALA 37

RESULT 975
Q988S7 RHIL0
ID Q988S7 RHIL0 PRELIMINARY; PRT; 77 AA.
AC Q988S7;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Mel6615 protein.
GN OrderedLocusNames=mel6615;
OS Rhizobium loti (Mesorhizobium loti).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OX Phyllobacteriaceae; Mesorhizobium.
 RN NCBI_TaxID=381;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Takeuchi Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 Mesorhizobium loti";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; BA000012; BAB52870.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 77 AA; 7826 MW; 6F6CC93AAA4A86BD CRC64;

Query Match 5.1%; Score 6; DB 2; Length 77;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 VLAALA 25
 |||||
 Db 7 VLAALA 12

RESULT 976

Q7YD61 RABIT
 ID Q7YD61_RABIT PRELIMINARY; PRT; 78 AA.
 AC Q7YD61;
 DT 01-OCT-2003 (TRENBLrel. 25, Created)
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
 DE Cytochrome b (Fragment).
 GN Name=cytb;
 OS Oryctolagus cuniculus (Rabbit).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
 OC Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Mougil F., Gautier A., Queney G., Sanchez M., Dennebouv N.,
 RA Monnerot M.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AJ539442; CAD62477.1; -; Genomic_DNA.
 DR SMR; Q7YD61; 1-78.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0046020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B C; 1.
 DR PROSITE; PS51003; CYTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 FT Respiratory chain; Transmembrane; Transport.
 FT NON_TER 1
 FT NON_TER 78
 SQ SEQUENCE 78 AA; 8726 MW; 87B36DC33B714D1B CRC64;

Query Match 5.1%; Score 6; DB 2; Length 78;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 LGGVLA 22
 |||||
 Db 3 LGGVLA 8

RESULT 977

Q56UF8 LYMST
 ID Q56UF8_LYMST PRELIMINARY; PRT; 78 AA.
 AC Q56UF8;
 DT 10-MAY-2005 (TRENBLrel. 30, Created)
 DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
 DE Nucleosome assembly protein-like protein (Fragment).
 OS Lymnaea stagnalis (Great pond snail)
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Basommatophora;
 OC Lymnaeidae; Lymnaeidae; Lymnaea.
 OX NCBI_TaxID=6523;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=CNS;
 RA Jaaro H., Levy Z., Fainzilber M.;
 RT "A genome-wide screening approach for membrane-targeted gene
 products";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY577332; AAT97090.1; -; mRNA.
 FT NON_TER 1
 FT NON_TER 79
 SQ SEQUENCE 79 AA; 8702 MW; 4A5304F100645477 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 78;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 GVLAAAL 24
 |||||
 Db 21 GVLAAAL 26

RESULT 978

Q56UF8 LYMST
 ID Q56UF8_LYMST PRELIMINARY; PRT; 79 AA.
 AC Q56UF8;
 DT 10-MAY-2005 (TRENBLrel. 30, Created)
 DT 10-MAY-2005 (TRENBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TRENBLrel. 30, Last annotation update)
 DE Nucleosome assembly protein-like protein (Fragment).
 OS Lymnaea stagnalis (Great pond snail)
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Basommatophora;
 OC Lymnaeidae; Lymnaeidae; Lymnaea.
 OX NCBI_TaxID=6523;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=CNS;
 RA Jaaro H., Levy Z., Fainzilber M.;
 RT "A genome-wide screening approach for membrane-targeted gene
 products";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY577332; AAT97090.1; -; mRNA.
 FT NON_TER 1
 FT NON_TER 79
 SQ SEQUENCE 79 AA; 8702 MW; 4A5304F100645477 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 1.3e+03; Mismatches 0; Indels 0; Gaps 0;

Qy 19 GVLAAL 24
 Db 18 GVLAAL 23

RESULT 979
 GCHI_MUCHA STANDARD; PRT; 80 AA.
 AC PS1598;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE GTP cyclohydrolase I (EC 3.5.4.16) (GTP-CH-I) (Fragment).
 OS Mucuna hirsuta.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 OC rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 OC Mucuna.
 NCBI_TaxID=40337;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Maier J., Witter K., Gietlich M., Ziegler I., Werner T., Ninnemann H.;
 RT "Homology cloning of GTP-cyclohydrolase I from various unrelated
 RT eukaryotes by reverse-transcription polymerase chain reaction using a
 RT general set of degenerate primers";
 RL Biochem. Biophys. Res. Commun. 212:705-711(1995).
 CC -1- CATALYTIC ACTIVITY: GTP + 2 H₂O = formate + 2-amino-4-hydroxy-6-
 CC (erythro-1,2,3-trihydroxypropyl)-dihydropteridine triphosphate.
 CC -1- ENZYME REGULATION: GTP shows a positive allosteric effect, and
 CC tetrahydrobiopterin inhibits the enzyme activity (By similarity).
 CC -1- PATHWAY: Cofactor biosynthesis; tetrahydrofolate biosynthesis; 2-
 CC amino-4-hydroxy-6-hydroxymethyl-7,8-dihydropteridine diphosphate
 CC from GTP: step 1.
 CC -1- SUBUNIT: Homopolymer (By similarity).
 CC -1- SIMILARITY: Belongs to the GTP cyclohydrolase I family.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; Z49759; CAA89829.1; -; mRNA.
 CC HSP; P22288; 1158.
 CC InterPro: IPR001474; GTP cyclohydrol.
 CC Pfam; PF01227; GTP cyclohydrol; 1.
 CC ProDom; PD003330; GTP cyclohydrol; 1.
 CC PROSITE; PS00859; GTP_CYCLOHYDROL.1.1; PARTIAL.
 CC PROSITE; PS00860; GTP_CYCLOHYDROL.1.2; 1.
 CC AllostERIC enzyme; Hydrolase; Tetrahydrobiopterin biosynthesis.
 CC DISULFID 4 75 By similarity.
 FT NON_TER 1 1
 FT NON_TER 80 80
 SQ SEQUENCE 80 AA; 8846 MW; F5F6C04D2BDDCC36 CRC64;
 Query Match 5.1%; Score 6; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 84 GKVLGL 89
 Db 23 GKVLGL 28

RESULT 980
 Q5XNS5_ANOGA
 ID Q5XNS5_ANOGA PRELIMINARY; PRT; 80 AA.
 AC Q5XNS5;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Glutathione-dependent peroxidase (Fragment).
 GN Name=GRX2;
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;
 OC Anophelinae; Anopheles.
 NCBI_TaxID=7165;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Ranson H., David J.P., Strode C., Vontas J., Nikou D.;
 RT "A microarray for studying metabolic resistance to insecticides in
 RT malaria vectors";
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY745230; AU93510.1; -; mRNA.
 DR GO; GO:0004601; P:peroxidase activity; IEA.
 DR InterPro; IPR010357; DUF953 thioredox.
 DR PANTHER; PTHR12452; DUF953; 1.
 DR Pfam; PF06110; DUF953; 1.
 KW Peroxidase.
 FT NON_TER 1 1
 SQ SEQUENCE 80 AA; 9282 MW; 8AF2DA39845AFE29 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 68 AAPYIE 73
 Db 2 AAPYIE 7

RESULT 981
 Q743K5_MYCPA
 ID Q743K5_MYCPA PRELIMINARY; PRT; 80 AA.
 AC Q743K5;
 DT 05-JUN-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=MAP0580C;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium avium complex (MAC).
 NCBI_TaxID=1770;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC SFRAIN-k10;
 RL Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017229; AAS02897.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 80 AA; 8491 MW; AE00014A77C6ABDB CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 Db 16 LAALAA 21

RESULT 982
 Q5ZB39_ORYSA
 ID Q5ZB39_ORYSA PRELIMINARY; PRT; 81 AA.
 AC Q5ZB39;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein BL151H08.9 (Hypothetical protein

DR B1045F02.50).
 GN Name=B1151H08.9; Synonyms=B1045F02.50;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Zhong H., Iwana H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 DR EMBL; AP003336; BAD53184.1; -: Genomic DNA.
 DR EMBL; AP003329; BAD53098.1; -: Genomic DNA.
 KW Hypothetical protein.
 SQ SEQUENCE 81 AA; 9139 MW; 4E4D5BA6C11B9833 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 17 LGGVLA 22
 Db 69 LGGVLA 74
 RESULT 983
 Q69N83 Oryza PRELIMINARY; PRT; 81 AA.
 ID Q69N83 Oryza PRELIMINARY; PRT; 81 AA.
 AC Q69N83 Oryza PRELIMINARY; PRT; 81 AA.
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Hypothetical protein OSUNBa0019B22.30 (Hypothetical protein
 DE P0450F09.3).
 GN Name=OSUNBa0019B22.30; Synonyms=P0450F09.3;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Katayose Y.;
 RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
 RT clone:OSUNBa0019B22.30";
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Sasaki T., Matsumoto T., Hattori M., Sakaki Y., Katayose Y.;
 RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, PAC
 RT clone:P0450F09.3";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005687; BAD36234.1; -: Genomic DNA.
 DR EMBL; AP005586; BAD36165.1; -: Genomic DNA.
 DR Gramene; Q69N83; -;
 KW Hypothetical protein.
 SQ SEQUENCE 81 AA; 8973 MW; 64EB829DA2926888 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 17 LGGVLA 22
 Db 69 LGGVLA 74
 RESULT 984
 Q06779 MYCTU PRELIMINARY; PRT; 81 AA.
 ID Q06779 MYCTU PRELIMINARY; PRT; 81 AA.
 AC Q06779 MYCTU PRELIMINARY; PRT; 81 AA.
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocustNames=Rv0660C;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
 RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
 RA Tekair F., Badcock K., Basham D., Brown D., Chillingworth T.,
 RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
 RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
 RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL; BX842574; CAB09386.1; -: Genomic DNA.
 DR PIR; G70534; G70534.
 DR TubercuList; Rv0660C; -;
 DR GO; GO:0003677; P:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 KW Complete proteome.
 SQ SEQUENCE 81 AA; 9169 MW; 29985341768D033D CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 21 LAALAA 26
 Db 37 LAALAA 42
 RESULT 985
 Q7ULI7 MYCBO PRELIMINARY; PRT; 81 AA.
 ID Q7ULI7 MYCBO PRELIMINARY; PRT; 81 AA.
 AC Q7ULI7 MYCBO PRELIMINARY; PRT; 81 AA.
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein Mb0679c.
 GN OrderedLocustNames=Mb0679c;
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1765;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
 RA Garnier T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,

RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsempe C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewison R.G.;
 RT "The complete genome sequence of *Mycobacterium bovis*."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
 DR EMBL; BX248336; CAD93541.1; -; Genomic_DNA.
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 81 AA; 9169 MW; 29985341768D033D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
 Db 37 LAALAA 42
 |||||

RESULT 986
 Q8Z0B9 ANASP PRELIMINARY; PRT; 81 AA.
 AC Q8Z0B9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Asr0179 protein.
 GN OrderedLocuNames=asr0179;
 OS Anabaena sp. (strain PCC 7120).
 OC Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]

NUCLEOTIDE SEQUENCE.
 RP MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium *Anabaena* sp. strain PCC 7120."
 RL DNA Res. 8:205-213(2001).
 DR EMBL; BA000019; BAB7703.1; -; Genomic_DNA.
 DR PIR; AC1829; AC1829.
 DR InterPro; IPR008538; DUF820.
 DR Pfam; PF05685; DUF820; 1.
 KW Complete proteome.
 SQ SEQUENCE 81 AA; 9141 MW; 8ABF95A6AFPF173D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 PAIVPD 51
 Db 66 PAIVPD 71
 |||||

RESULT 987
 Q98MZ8 RHIL0 PRELIMINARY; PRT; 81 AA.
 AC Q98MZ8;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Mar0370 protein.
 GN OrderedLocuNames=mar0370;
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]

RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=MAFF303099;
 RA MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT *Mesorhizobium loti*."
 RL DNA Res. 7:331-338(2000).
 DR EMBL; BA000012; BAB47965.1; -; Genomic_DNA.
 KW Complete proteome.
 SQ SEQUENCE 81 AA; 8759 MW; 519182D79F75698C CRC64;

Query Match 5.1%; Score 6; DB 2; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 AALAA 27
 Db 47 AALAA 52
 |||||

RESULT 988
 O96694 SAVES PRELIMINARY; PRT; 82 AA.
 AC O96694;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Cytochrome b (fragment).
 OS *Cuculus canorus* (common cuckoo).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Cuculiformes; Cuculidae; Cuculus.
 OX NCBI_TaxID=55661;
 RN [1]

NUCLEOTIDE SEQUENCE.
 RP TISSUE=Blood;
 RA Aragon S., Moller A.P., Soler J.J., Soler M.;
 RT "Molecular phylogeny of cuckoos supports a polyphyletic origin of
 RT brood parasitism."
 RL J. Evol. Biol. 12:495-506(1999).
 CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
 CC cytochrome c1 and the Rieske protein (By similarity).
 DR EMBL; AF072611; AAC70973.1; -; Genomic_DNA.
 DR SMR; O96694; 1-82.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005746; C:mitochondrial electron transport chain; IEA.
 DR GO; GO:0005739; C:mitochondrion; IEA.
 DR GO; GO:0046872; F:metal ion binding; IEA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR005798; Cytb_b6_C.
 DR Pfam; PF00032; Cytochrom B C; 1.
 DR PROSITE; PS1003; CVTB_CTER; 1.
 KW Electron transport; Heme; Iron; Metal-binding; Mitochondrion;
 KW Respiratory chain; Transmembrane; Transport.
 FT NON TER 1 82
 FT NON TER 1 82

Query Match 5.1%; Score 6; DB 2; Length 82;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 LGGVLA 22
 Db 18 LGGVLA 23
 |||||


```

RESULT 989
O99697 GEOCA
ID O99697 GEOCA PRELIMINARY; PRT; 82 AA.
AC O99697;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Cytochrome b (fragment).
OS Geococcyx californianus (Roadrunner).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Cuculiformes; Neomorphidae; Geococcyx.
OX NCBI_TaxID=8947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Feather;
RA Aragon S., Moller A.P., Soler J.J., Soler M.;
RT "Molecular phylogeny of cuckoos supports a polyphyletic origin of
RT brood parasitism.";
RL J. Evol. Biol. 12:495-506(1999).
CC -1- SUBUNIT: The main subunits of complex b-c1 are: cytochrome b,
CC cytochrome c1 and the Rieske protein (By similarity).
DR EMBL: AF072614; AAC70976.1; -; Genomic_DNA.
DR SMR: O99697; 1-82.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005746; C:mitochondrial electron transport chain; IEA.
DR GO: GO:0005739; C:mitochondrion; IEA.
DR GO: GO:0046872; F:metal ion binding; IEA.
DR GO: GO:0016491; F:oxidoreductase activity; IEA.
DR GO: GO:0006118; F:electron transport; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR005798; Cytb_b6_C.
DR Pfam: PF00032; Cytochrom B C; 1.
DR PROSITE: PS51003; CYTB_CTEF; 1.
KW Electron transport; Heme; iron; Metal-binding; Mitochondrion;
KW Respiratory chain; Transmembrane; Transport.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9295 MW; 87703E1469EC3E30 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 82;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 18 LGGVLA 23

RESULT 990
O9HRW2 HALSA
ID O9HRW2 HALSA PRELIMINARY; PRT; 83 AA.
AC O9HRW2;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Vng0516h.
GN OrderedlocusNames=VNG0516H;
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=NRC-1 / ATCC 700922 / JCM 11081;
RX MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorason V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J.J., Dahl T.A., Weltri R., Goo Y.A.,
RA Leichauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,

```

```

RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL: AE005003; AAG19046.1; -; Genomic_DNA.
DR PIR: B84210; B84210.
KW Complete proteome.
SQ SEQUENCE 83 AA; 8479 MW; FC3FD785E802B4CE CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 31 LGGVLA 36

RESULT 991
O69KP4 ORYSA
ID O69KP4 ORYSA PRELIMINARY; PRT; 83 AA.
AC O69KP4;
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Hypothetical protein OSJNBa0064123.7.
GN Name=OSJNBa0064123.7;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Sasaki T., Matsumoto T., Katayose Y.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC
RT clone:OSJNBa0064123.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005912; BAD36496.1; -; Genomic_DNA.
DR Gramene: O69KP4; -.
KW Hypothetical protein.
SQ SEQUENCE 83 AA; 9303 MW; B53D2F69F3C3CFE CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 LGGVLA 22
Db 71 LGGVLA 76

RESULT 992
O84QX9 ORYSA
ID O84QX9 ORYSA PRELIMINARY; PRT; 83 AA.
AC O84QX9;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein OSJNBa0093113.17.
GN Name=OSJNBa0093113.17;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.B., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Fadrosch D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feildblyum T.V.,

```

RA Yang Q.O., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Buell R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC097279; AA04199.1; -; Genomic_DNA.
 DR Gramene; Q84QX9; -;
 DR InterPro; IPR002345; Lipocalin.
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 83 AA; 8799 MW; 2C575580633A46FA CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAA 26
 DB 14 LAALAA 19
 |||||
 RESULT 993
 Q8VKG8 MYCTU
 ID Q8VKG8 MYCTU PRELIMINARY; PRT; 83 AA.
 AC Q8VKG8;
 DT 01-MAR-2002 (TEMBLrel. 20, Created)
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=MT0689;
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OC Mycobacterium tuberculosis complex.
 OX NCBI_TaxID=1773;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=CDC 1551 / Oshkosh;
 RX MEDLINE=22206494; PubMed=12218036;
 RX DOI=10.1128/JB.184.19.5479-5490.2002;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwin M.L., Haft D.H.,
 RA Hickey E.K., Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
 RA Salzberg S.L., Delcher A., Uterback T.R., Weidman J.F., Khouri H.M.,
 RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
 RA Fraser C.M.;
 RA "Whole-genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL J. Bacteriol. 184:5479-5490 (2002).
 DR EMBL; AE000516; AAK44914.1; -; Genomic_DNA.
 DR TIGR; MT0689; -;
 KW Hypothetical protein.
 SQ SEQUENCE 83 AA; 9401 MW; 04F447485BA90A7C CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAA 26
 DB 39 LAALAA 44
 |||||
 RESULT 994
 Q7VVF9 BORPE
 ID Q7VVF9 BORPE PRELIMINARY; PRT; 83 AA.
 AC Q7VVF9;
 DT 01-OCT-2003 (TEMBLrel. 25, Created)
 DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE Molybdopterin converting factor.

GN Name=moad; OrderedLocusNames=BP2710;
 OS Bordetella pertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=520;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=tohami I / ATCC BAA-589 / NCTC 13251;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
 RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jegerle K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RA "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica";
 RL Nat. Genet. 35:32-40 (2003).
 DR EMBL; BX640419; CAE42987.1; -; Genomic_DNA.
 DR HSSP; P30748; IJW9.
 DR GO; GO:0006790; P:sulfur metabolism; IEA.
 DR InterPro; IPR003749; This.
 DR Pfam; PF02597; This; 1.
 KW Complete proteome.
 SQ SEQUENCE 83 AA; 8831 MW; APD53E0DE5CB9EB1 CRC64;
 Query Match 5.1%; Score 6; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 1.3e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 21 LAALAA 26
 DB 37 LAALAA 42
 |||||
 RESULT 995
 Q7WB93 BORPA
 ID Q7WB93 BORPA PRELIMINARY; PRT; 83 AA.
 AC Q7WB93;
 DT 01-OCT-2003 (TEMBLrel. 25, Created)
 DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
 DE Molybdopterin converting factor.
 GN Name=moad; OrderedLocusNames=BP1113;
 OS Bordetella parapertussis.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Alcaligenaceae; Bordetella.
 OX NCBI_TaxID=519;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=12822 / ATCC BAA-587;
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngl1227;
 RA Parkhill J., Sebatia M., Preston A., Murphy L.D., Thomson N.R.,
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
 RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jegerle K.,
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
 RA "Comparative analysis of the genome sequences of Bordetella pertussis,
 RT Bordetella parapertussis and Bordetella bronchiseptica";
 RL Nat. Genet. 35:32-40 (2003).
 DR EMBL; BX640426; CAE36414.1; -; Genomic_DNA.
 DR HSSP; P30748; IJW9.
 DR GO; GO:0006790; P:sulfur metabolism; IEA.
 DR InterPro; IPR003749; This.

DR Pfam: PF02597; This: 1.
KW Complete proteome.
SQ SEQUENCE 83 AA; 8829 MW; 72D4DE0DE5CB9EA3 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 LAALAA 26
| | | | |
DB 37 LAALAA 42

RESULT 996

Q7WMR1_BORBR PRELIMINARY; PRT; 83 AA.
AC Q7WMR1

DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Molybdopterin converting factor.
GN Name=moaD; OrderedLocusNames=BE1329;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsis K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinowitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.;
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640441; CAE31827.1; -: Genomic_DNA.
DR HSSP; P30748; 1JW9.
DR GO; GO:0006790; P:sulfur metabolism; IEA.
DR InterPro; IPR003749; This.
DR Pfam; PF02597; This; 1.
KW Complete proteome.

SQ SEQUENCE 83 AA; 8815 MW; BID4DE0DE5CB9EB1 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 LAALAA 26
| | | | |
DB 37 LAALAA 42

RESULT 997

Q9A171_STRPY PRELIMINARY; PRT; 83 AA.
AC Q9A171

DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SPY0432;
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;

[1] NUCLEOTIDE SEQUENCE.

RN STRAIN=SF370 / ATCC 700294 / Serotype M1;
RC MEDLINE=21192684; PubMed=11296296; DOI=10.1073/pnas.071559398;
RA Perretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006504; AAK33453.1; -: Genomic_DNA.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 83 AA; 9123 MW; 017DCDB8685A5419 CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 20 VLAALA 25
| | | | |
DB 62 VLAALA 67

RESULT 998

Q9S2U4_STRCO PRELIMINARY; PRT; 83 AA.
AC Q9S2U4

DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative membrane protein.
GN OrderedLocusNames=SCO2041; ORFNames=SC4G6.10c;
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]

NUCLEOTIDE SEQUENCE.

RC STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953; DOI=10.1038/417141a;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L.D., Oliver K., O'Neill S.,
RA Rabinowitch E., Rajandream M.A., Rutherford K.M., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J.R., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL939111; CAB51433.1; -: Genomic_DNA.
DR FIR; T35070; T35070.
KW Complete proteome.

SQ SEQUENCE 83 AA; 8567 MW; 286619B1095EC83D CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 VLLGGV 20
| | | | |
DB 45 VLLGGV 50

RESULT 999

Q8P292_STRP8 PRELIMINARY; PRT; 83 AA.
AC Q8P292

DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)

Search completed: January 27, 2006, 19:33:11
Job time : 219 secs

```

DE Hypothetical protein spyM18_0479.
GN OrderedLocusNames=spyM18_0479;
OS Streptococcus pyogenes (serotype M18).
SC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
NCBI_TaxID=301451;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108; DOI=10.1073/pnas.062526099;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
DR EMBL; AE009987; AAL97203.1; -, Genomic_DNA.
KW Complete proteome.
SQ SEQUENCE 83 AA; 9131 MW; B06BA754B835A98F CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 VLAALA 25
   |||||
Db 62 VLAALA 67

RESULT 1000
Q608F1_METCA
ID Q608F1_METCA PRELIMINARY; PRT; 83 AA.
AC Q608F1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Conserved domain protein.
GN OrderedLocusNames=MCA1541;
OS Methylococcus capsulatus.
SC Bacteria; Proteobacteria; Gammaproteobacteria; Methylococcales;
OC Methylococcaceae; Methylococcus.
NCBI_TaxID=414;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Bath / NCIMB 11132;
RX PubMed=15383840; DOI=10.1371/journal.pbio.0020303;
RA Ward N.L., Larsen O., Sakwa J., Bruseeth L., Khouri H.M., Durkin A.S.,
RA Dimitrov G., Jiang L., Scanlan D., Kang K.H., Lewis M.R., Nelson K.E.,
RA Mathe B.A., Wu M., Heidelberg J.F., Paulsen I.T., Fouts D.E.,
RA Ravel J., Tettelin H., Ren Q., Read T.D., DeBoy R.T., Seshadri R.,
RA Salzberg S.L., Jensen H.B., Birkeland N.K., Nelson W.C., Dodson R.J.,
RA Grindhaug S.H., Holt I.E., Bidhammer I., Jonassen I., Vanaken S.,
RA Uterback T.R., Feldblyum T.V., Fraser C.M., Lillehaug J.R.,
RA Eissen J.A.;
RT "Genomic insights into methanotrophy: the complete genome sequence of
RT Methylococcus capsulatus (Bath).",
RL PLoS Biol. 2:1616-1628(2004).
DR EMBL; AE017282; AAU92212.1; -, Genomic_DNA.
DR TIGR; MCA1541; -.
DR InterPro; IPR005631; DUF339.
DR Pfam; PF03937; TPR_div1; 1.
KW Complete proteome.
SQ SEQUENCE 83 AA; 9242 MW; 9A27C7260CD208BF CRC64;

Query Match 5.1%; Score 6; DB 2; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 LAALAA 26
   |||||
Db 67 LAALAA 72

```